omega

The International Journal of Management Science

Volume 5 1977

Author and Title Index

● **Chief Editor**

Professor SAMUEL EILON, DSc(Eng.), PhD — Department of Management Science, Imperial College, London, SW7 2BX

● **Associate Editors**

RJ BETTS and RB FLAVELL — Department of Management Science, Imperial College, London, SW7 2BX

Editorial Board

Professor RJ BALL, London Graduate School of Business Studies, Sussex Place, Regent's Park, London NW1 4SA

Professor BEN BERNHOLTZ, Department of Industrial Engineering, University of Toronto, Toronto 5

Professor C WEST CHURCHMAN, Space Sciences Laboratory, University of California, Berkeley, CA 94720

Professor WW COOPER, Graduate School of Business Administration, Harvard University, Boston, MA 02163

Professor BELA GOLD, Department of Economics, Case Western Reserve University, Cleveland, OH 44106

Professor CH GUDNASON, Danmarks Tekniske Højskole, København

Professor DOUGLAS C HAGUE, Manchester Business School, Booth Street West, Manchester M15 6PB

DAVID B HERTZ, McKinsey and Company Inc., 245 Park Avenue, New York, NY 10017

Professor KNUT HOLT, The Norwegian Institute of Technology, 7034 Trondheim

Professor HAROLD J LEAVITT, Graduate School of Business, Stanford University, Stanford, CA 94305

Professor GERALD J LIEBERMAN, Department of Operations Research, Stanford University, Stanford CA 94305

Sir FRANK MCFADZEAN, British Airways, PO Box 13, Air Terminal, Buckingham Palace Road, London SW1W 9SR

Professor DANIEL TEICHROEW, Department of Industrial and Operations Engineering, University of Michigan, Ann Arbor, MI 48104

Pergamon Press Oxford · New York · Paris · Frankfurt

AUTHOR INDEX

Ackoff RL 649
Baker C 117
Balachandran KR 185
Becker PE 87
Bedi DN 721
Bennett PG 749
Betts RJ 381
Brown RM 746
Browne WG 7
Burton RM 395
Carlson PG 271
Carlson RC 57
Chapman Findlay III 93
Clark AW 509
Colcutt RH 117
Cole HSD 529
Colombo U 511
Cosmetatos PG 673
Cucuy S 333
Dando MR 89, 255, 506
Dar-El EM 333
Davis W 709
Defrenne A 89
Dunbar RLM 309
Eden C 255
Eilon S 1, 107, 227, 361, 437, 499, 627, 673
Ericson A 309
Finn R 133, 557
Fitts RL 207
Flavell RB 543
Flowers AD 473
Ford H 121
Furtan H 43
Gallagher CC 731
Gentzler GL 215
Gerstenfeld A 35
Gleason JM 611
Goldberg WH 309
Gowens JW 609
Graham M 317
Harari E 161
Harper DG 239
Hershauer JC 609
Higgins JC 133, 557
Howells M 317
Huang CC 309
Hull JC 605
Jagetia LC 741
Jenkins JO 115, 317
Jones HG 583, 689
Joynt P 425
Keown AJ 618
Khalil TM 215
Kynaston Reeves T 239
Ladany SP 721
Laulajainen R 67
Leonidis-Plessas D 7
Lev B 193
Lietaer BA 149
Lingaraj BP 348
Lopez DA 614
Lowe EA 173
Mace JR 567
Mason JB 207
Mathewson SC 7
Mathusz DV 593
Matthews JP 23
McCarl BA 43
Meade N 663
Misra RB 99
Moerland PW 343
Morecroft JDW 699
Moskowitz H 43
Obel B 395
Plotkin HA 77
Preston SE 473
Remus W 349
Rivett BHP 367, 505
Rothwell R 415
Salomon J-J 635
Sarin RK 481
Sharp RG 89, 506
Sivazlian BD 215
Soyster AL 193
Spiller P 415
Stanbury W 309
Talavage J 709
Taylor BW III 618
Teubal M 415
Therp HH 57
Thiétart RA 281
Thompson ER 93
Thompson HE 23
Tinker AM 173
Tomkins R 235
Toof DI 193
Townsend J 415
Turner RE 11
Uri ND 463
Van Gigch JP 9
Van Slyke WJ 614
Vertinsky I 309
Waters CDJ 317
Walton JHD 293
Watson-Gandy CDT 699
Winckles K 115
Wolfe H 348
Wortham AW 99
Zeira Y 161

Vol. 5
1977

TITLE INDEX

Accuracy in the Long Run—Where Are We Now? 529
Analysis of a Partially-Observed Queuing System 77
Chief Executive and his Information System 557
Computer-Based Scheduling of Technical Resources 317
Conservative Linear Programming with Mixed Multiple Objectives 193
Contingency Analysis for Effective Administration 425
Contribution of the Social Sciences 635
Criteria for the Selection of Reporting Method 567
Decline of Venture Capital 663
Divisionalization and Transfer Pricing: A Review 543
Dynamic Booking Rules for Flights with an Intermediate Stop 721
Employee Attitude Surveys: Their Interpretation and Use 239
Forecasting: A Hybrid Approach 463
How Do You Feel About Your Work? (Editorial) 227
How to Charge for Management Services (Feedback) 115, 235
Human Dynamics: A New Approach to Human Behavior 281
Innovation in a Waste Economy 121
Insurance Exposure and Investment Risks: Management and Regulatory Implications 23
Interactive LP Models in Management—An Analogue Approach 293

Interdependence and Innovation 35
International Register of Theses and Dissertations 101, 221, 353, 491, 621, 753
Investment and the Swedish Economy 583
Judgmental Measures for Public Project Selection 11
Maintaining the Calibration of an Information System 271
Management of Timber and Related Industries in Sweden 689
Management Perspectives in Physical Distribution 437
Management Science Approach to Planning Freedom (Feedback) 9
Market Representation in the Motor Industry 699
Marketing Segmentation Research—An Application to Bank Services 207
Memoranda 87, 343, 605, 741
Method for Short-Term Profit Analysis 309
More Against Optimization (Editorial) 627
Multicriteria Approach to Strategic Planning: An Application in Inventory Control 57
Multidimension Scaling for Multiobjective Policies 367
Multilevel Approach to Organizational Issues of the Firm—A Critical Review 395
New Directions for Management Accounting 173
Optimal Design of a Post Office 185
Optimal Mixed-Model Sequencing for Balanced Assembly Lines 333
Planning for Freedom (Feedback) 117
Planning Model for Underground Mines—An Application in Developing Country 149
Planning Models in the UK: A Survey 133
Production Technology and Management in Planned and Market Economies 731
Profitability Model for Tactical Planning 673
Quadratic Programming Applications 43
Quality of Working Life—A Comment (Feedback) 509
Quantitative Models for Optimal Rest Period Scheduling 215
Reflections on Operational Research. A Report from the Euro II Congress 255
Review of Inflation Accounting and its Economic Implications 381
Scheduled and Charter Impairment in North Atlantic Air Services (Feedback) 7
School for Public Policy (Editorial) 361
Screening of Multiattribute Alternatives 481
Some Methodological Aspects of Innovation Research 415
Spatial Economies in Distribution Transports 67
Strategies for Europe: Proposals for Science and Technology Policies—Industrial Innovation in Europe 511
Structural Determinism (Editorial) 499
Structural Mapping for Policy Selection (Feedback) 505
Structural Sources of Personnel Problems in Multinational Corporations: Third-Country Nationals 161
Technician or Adviser? (Editorial) 1
Technicians, Advisers—and All-Rounders in Operational Research? (Feedback) 506
Three-Level Models for Hierarchical Coordination 709
Towards Flexible Organizations: A Multidimensional Design 649
Value of Information Concept Applied to Data Systems 593
Who Says Production Control Is Dead? (Editorial) 107
Work Force Scheduling with the Search Decision Rule 473

1. 5
77

Aims and Scope

The purpose of OMEGA is to provide a vehicle for reporting developments in management science, operational research and managerial economics, including research results and applications. Published material ranges from original contributions to review articles describing the state of the art in specific areas, together with shorter critical assessments of particular management techniques. Other features of the journal are the Memoranda section for short communications, Feedback (a correspondence column) and a register for theses and dissertations.

The journal is intended to serve the needs of practising managers, specialists in management services, operational research workers and management scientists, management consultants, academics, students and research personnel.

Annual Subscription Rates (1978)

Published bi-monthly. For libraries, university departments, government laboratories, industrial and other multiple-reader institutions U.S. $105.00 per annum (including postage and insurance). *Specially reduced rates to individuals:* In the interests of maximising the dissemination of material published in this important international journal we have established a two-tier price structure. Any individual, whose institution takes out a library subscription, may purchase a second or additional subscription for their use at the much reduced rate of U.S.$30.00 (£12.50) *per annum*. All subscription enquiries should be addressed to: The Subscription Fulfillment Manager, Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, England.

Microform Subscriptions and Back Issues

Back issues of all previously published volumes are available in the regular editions and on microfilm and microfiche. Current subscriptions are available on microfiche simultaneously with the paper edition and on microfilm on completion of the annual index at the end of the subscription year.

Publishing, Subscription and Advertising Offices

Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, England.
Pergamon Press Inc., Maxwell House, Fairview Park, Elmsford,
New York 10523, U.S.A.

Copyright 1977 Pergamon Press Ltd.

No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means: electronic, electrostatic, magnetic tape, mechanical, photocopying, recording or otherwise, without permission in writing from the publishers.

Printed in Great Britain by A. Wheaton & Co., Exeter

OMEGA

The International Journal of Management Science

Volume 5, Number 1, 1977

Contents

SAMUEL EILON	1	Editorial—Technician or Adviser?
	7	Feedback
RONALD E TURNER	11	Judgmental Measures for Public Project Selection
JOHN P MATTHEWS, HOWARD E THOMPSON	23	Insurance Exposure and Investment Risks: Management and Regulatory Implications
A GERSTENFELD	35	Interdependence and Innovation
BRUCE A MCCARL, HERBERT MOSKOWITZ, HARTLEY FURTAN	43	Quadratic Programming Applications
ROBERT C CARLSON, HANS H THERP	57	A Multicriteria Approach to Strategic Planning: An Application in Inventory Control
R LAULAJAINEN	67	Spatial Economies in Distribution Transports
HOWARD A PLOTKIN	77	Analysis of a Partially-Observed Queuing System
PE BECKER	87	Memoranda
MR DANDO, A DEFRENNE, RG SHARP	89	
M CHAPMAN FINDLAY III, ER THOMPSON	93	
RB MISRA, AW WORTHAM	98	
—	101	International Register of Theses and Dissertations

BLANK
PAGE

Vol. 5
1977

OMEGA The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

EDITORIAL

Technician or Adviser?

THERE IS a semi-apocryphal account that I have recently come across, of a business enterprise in which a great deal of thought had been given to the question of how to organize its OR (Operational Research) department, and how to determine the scope of its work. The account contains many over-simplifications with respect to detail, but its portrayal of the problem is essentially accurate and is reminiscent of similar cases to be found elsewhere. Being a relatively large enterprise, in which the use of computers has mushroomed beyond all expectations in a single decade, the transformation of administrative tasks and the complete overhaul of the communication network necessitated a mammoth exercise involving systems analysts and MIS (Management Information System) experts of every conceivable denomination. At first large computerized systems in the business environment appear to liberate the organization from its traditional shackles and allow administrative procedures to be redesigned, but—as one soon learns—it is naïve to expect that after a flurry of activity at the implementation stage, and after the computers have passed their prescribed operational tests, the systems can be left to perform their duties with only a skeleton staff of maintenance men. In reality, the computerized system demonstrates a seemingly insatiable appetite for personnel to attend to its needs: as soon as a complex control procedure is implemented, new requirements arise, either as a result of exogenous forces or through changing internal circumstances, and the constraints imposed on the enterprise by newly implemented procedures then become only too apparent. With all the care and ingenuity that systems analysts bring to bear on the design of new systems, there is a limit to the level of flexibility that can be incorporated in such designs, and it is clearly impossible to anticipate all the future requirements or new circumstances that the systems would have to cope with.

It therefore follows, and this conclusion is borne out by observation, that a large influx of professional manpower at the installation phase is not a passing phenomenon, and that (except for installation engineers) a continuous presence of management services staff is required to maintain implemented systems, to monitor their performance, to update or replace them and to integrate and adapt their use to the changing needs of the organization.

In the enterprise that I referred to, this continuous role is fully recognized, and its relatively large management services department is almost entirely devoted to designing, implementing, operating and maintaining computer systems. The department is charged with the responsibility for producing a wide variety of printouts for the company's operations on a regular basis, and this routine work occupies the major part of the computer's time, but the department's responsibility also extends to writing and running *ad hoc* computer programs at the request of any authorized person in the organization.

This description gives a picture of a management services department largely concerned with well prescribed tasks; the routine nature of many of these tasks is often manifest by their relatively high frequency, but even when this is not the case, they are typical by their replication, namely by their common characteristics and by the fact that they require a well-defined set of tools and skills for their execution. This is not to say that the department has no *ad hoc* problems to tackle; every newly designed system needs to cater for the kind of information that various echelons in the hierarchy should have and the control functions that should be exercised. And since the designers of a new system are naturally anxious not to spend their efforts on merely computerizing an outdated manual set of procedures, but to use the opportunity to modernize the relevant administrative structure, it follows that they inevitably assume the role of change-agents and need to get involved in such questions as aims and means, at least in order to ascertain that their systems do indeed have the potential to perform in the way envisaged. Thus, the management services department is not totally deprived of opportunities to engage in innovative and imaginative tasks, but these come about as a direct result of the main brief given to the department, which is to computerize the appropriate data processing and information flows in the enterprise.

Vol. 5
1977

It is against this background, and the advent of a major restructuring exercise for the whole organization, that the company had to consider the position and role of OR. In the light of the main preoccupation of the management services department with computer systems, it was decided that projects involving investigations of existing or prospective operations should be undertaken within a separate planning department, which apart from corporate planning groups (one for short, and one for long term planning, a somewhat doubtful demarcation) has an operational research section (ORS) and a special assignments section (SAS). The placing of these sections within the planning department reflects the desire of the senior executives in the organization to ensure that results from any investigations undertaken by the two sections are taken into account in planning and budgeting, so that the operating divisions can be made to realize that the results are taken seriously and that certain expectations emerge from the implementation of approved recommendations of the investigatory teams.

But why two separate sections? Why not a single super department (by

whatever name is thought appropriate) with wide ranging expertise and capabilities to tackle a portfolio of projects of a diversified nature? Now, there are many ways in which one can advance arguments to support the rationale of two sections rather than one: for example, it may be administratively convenient to separate short term projects from long term ones, or projects which are tactical in nature from those concerned with strategy, or projects concerned with current operations from those relating to new ventures, or projects which aim to improve production and marketing activities from those involving problems of manpower and organizational design. Above all, the size of the department may well be a relevant factor; as the number of professionals in the department increases, there are clearly some advantages of scale, the analysts have a greater opportunity to gain experience from a wider variety of assignments and to learn from each other, there is more flexibility in the design of a portfolio or projects and greater toleration of experimental and exploratory work. But against these advantages of scale one must have regard to the increasing loss of individual identity at the lower echelons of the hierarchy and to the possible deterioration in the quality of both technical supervision and man-management in a larger group. There is, of course, no optimal size for OR groups which is accepted as universally valid, and what is regarded as a perfect size in one instance, may be thought of as unwieldy in another. But in the example that I have cited, many people could argue that two groups of thirty professionals each would be better than one of sixty.

What is so curious about the company in question is that none of these arguments is being advanced, at least not explicitly, as a justification for having two separate sections, although when you begin to question the rationale more closely, you find that some of the arguments are necessarily implied in the adopted organizational structure. The official position is that the ORS consists of people with a scientific background and knowledge of analytical tools, capable of constructing OR models (often necessitating the use of the computer); most of its recruits tend to be relatively young and many come straight from a university postgraduate course in OR, they all have demonstrable intellectual abilities, but not necessarily an intimate knowledge of the industry in which the company operates. In contrast, the SAS does not recruit from the outside and its members have no OR background or qualifications, but their proven ability and prior experience as bright junior or middle managers gives them a good understanding of the operational environment, and they can speak with some authority about the problems of the company and its competitors. Thus, the ORS has the technical expertise, while the SAS has the industrial and managerial expertise. Members of the ORS would not immediately comprehend the ramifications of a given operational problem and would need to engage in a fair amount of probing before they could manage to dispel the suspicions of the operating man; members of the SAS would not recognise a linear programming model if they saw one, come to think of it they would not be interested either.

Ostensibly, the logic behind this recruiting policy is to buy expertise in the most effective way: since OR men can scarcely be found internally, they are brought in from the outside, and since they do not grow on trees, it is logical to recruit them fairly young, and to put the emphasis on a sound technical background rather than on industrial experience; the SAS men, on the other hand, can be found internally, and as promising men in their thirties see a spell in the SAS as an extremely valuable step on their career ladder, there is no shortage of candidates to choose from. Thus, the rationale is complete, the ORS and the SAS complement each other, and together they provide both technical skills and business acumen.

If that were the whole story, we would only need to conclude with a note of congratulation to the company on its ability to resolve an intriguing personnel and organizational issue. But in managerial affairs a solution to one problem creates another and in every situation we have to determine whether the disease is less painful or more tolerable than the cure. In the case in point the natural question that springs to mind is: how does it work? In this context a simple test is to ask what happens to a new project and which section it is allocated to. Now, if the answer were that a project team is set up with personnel from both sections, where the composition of the team reflects the skills appropriate to that project (indeed personnel may be drawn from other parts of the company), then doubts about the organizational aspect of having two separate sections may somewhat dissipate. But the actual answer given in our case is that the project is assessed in terms of its content and with regard to the preponderance of skill required to tackle it: if it is clearly numerical and quantitative in character, if it involves the analysis of data, if it calls for the use of OR and statistical techniques, then it is assigned to the ORS, otherwise (and particularly when intimate knowledge of the workings of the organization is called for) it is given to the SAS.

Vol. 5
1977

Now, this sounds like a sensible answer; it recognises the advantage of clearly assigning total responsibility for the project to one section, thereby emphasizing the accountability of its management, and it takes cognizance of the fact that the two sections have different potential contributions to offer. But there are several features in this policy which are rather disquieting.

First, how do you determine in advance what methods of analysis would be appropriate for a given project? Is it conceivable that a project carried out in, or on behalf of, a planning department could be totally devoid of any numerical content? The current policy of assigning projects is therefore tantamount to prejudging the way in which they are to be carried out. Secondly, if the view is taken that OR can only be applied when a great deal of data is on record, then clearly the work of the ORS must largely be confined to the analysis of current operations, whereas SAS would tend to look at new ventures and products for which operational data are not available. Thirdly, major studies which involve intangibles often correspond to strategic problems, and would tend to go to the

SAS, whereas the ORS would gradually become more preoccupied with problems of a tactical nature. Furthermore, problems which require a systematic examination of processes which incorporate aspects of managerial control (and it is difficult to conceive of processes that do not) may well be thought of as the province of MIS, and in the context of the enterprise cited here, it would be natural for the management services department to press for ownership of any project that is not perceived as a direct input to the corporate planning function. Thus, the ORS faces 'competition' from both the SAS and MIS, each having a seemingly legitimate claim to sizeable chunks of project-work.

The sum total of this account is that the ORS may increasingly be regarded as a group of competent technicians, as the boys you call in to handle well-defined and well-prescribed problems, as the clever plumbers with their clever set of tools. This impression is naturally enhanced by the evident disparity between the average age as well as the operating experience of members of the ORS and SAS.

Many an OR man will be content with the role of a technician; it is undoubtedly an important and honourable role to play, and technicians rightly take immense pride in demonstrating whatever unique skills they have. But for some of us this is not enough. If OR is to make a significant impact on the decision-making process, and if it is to serve truly as a change-agent, then it should not be confined to doing a bit of computing on tactical and often short-term problems.

It would appear that there is a distinction to be made between the role of the technician and the role of the adviser. The former is given a problem and he sets about solving it, selecting whatever tool is suitable from his tool-kit, and if no suitable tool is readily available, he will proceed to adapt an existing tool or to devise a new one for the occasion. Having completed his work, he can often point to tangible benefits that may result, and he recedes gracefully from the scene with only a modicum of concern about implementing his ideas (he is, of course, anxious to have his recommendations implemented, but often begrudges the amount of time and effort that he would personally have to invest in teaching others what to him is perhaps patently obvious). The adviser, on the other hand, is not interested in a particular problem, but in how it has come about, whether valid questions are being asked, whether any prescribed constraints should be challenged; above all, the adviser is the man who helps to formulate the problem in the first instance, in defining objectives and in assessing wide ranging implications of given trends or actions on parts and on the whole of the enterprise. The adviser is a man who does not shrink from considering intangibles and who will use any method (numerical or otherwise) to face problems for which decisions have to be made. But the major difference between the technician and the adviser lies in their perception of the level of responsibility attached to their work. The former accepts only limited responsibility for the definition of problems and for their stated objectives, but total responsibility

for the execution of his projects, once they are clearly defined; the latter identifies himself with the decision process and sees it as his duty to give recommendations on courses of action which incorporate his assessment of objectives and values.

This distinction was brought home to me the other day when, after a long discussion about the results of an investigation conducted by several analysts on the activities of an operating division, the company director closed his file and—wagging his forefinger at me—said, "now let's forget about all these figures-men, and turn to the main issues". I hope that I am not alone in my conviction that forward looking OR men should indeed concern themselves with the main issues.

SAMUEL EILON
Chief Editor

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

FEEDBACK

HAVING read the paper on Scheduled and Charter Impairment in North Atlantic Air Services by W. G. Browne, we would like to make the following comments.

From the 1971 results, moving to the optimal point is shown to involve decreasing the foregone revenue of the charter carrier by $139 million and increasing the foregone revenue of the scheduled carriers by $114 million. But in 1971 the lowest scheduled fare was lower than the average charter fare ($105 vs. $115).

In the years 1972–1975 the situation was reversed and the average charter fare was cheaper than the cheapest scheduled fare. In this case, accepting all the author's assumptions, the actual traffic divisions must coincide with the optimum mix point. Thus 1971 is the only year where the model appears of value.

In any case surely it is trivial that the best revenue situation must exist where prices are pushed up in a market whose propensity to travel, as in the would-be-scheduled passengers, is independent of price. (As an aside, are the author's assumptions consistent? He assumes that a 20% increase in fares will cause a 60% attrition in the volume of charter passengers, yet supposes that a 10% increase for the scheduled traveller at the lowest end of the market will cause no loss of sales.)

An additional point in his method is the assumption that the fare structure in the indifference curve is a continuum (Fig. 5). Here his geometrically derived optimum implies a shift of 1 million while reference to Table 1 shows that 1.536 million passengers use the cheapest scheduled fare. Is this also the reason for slight inconsistencies in the figures quoted for his optimum shift? A foregone charter revenue of $139M implies 1.2M passengers at $115 each. As we are operating to the left of point A in his Fig. 4 a scheduled fare of $95 is implied and this is lower than any previously quoted fare.

Possibly, if the costs of providing a service can be estimated, the revenue foregone minimisation can be replaced by the notion of a total system profit foregone minimisation as an economic objective for defining the optimum mix. But to avoid trivial answers, the notion of price elasticity must be incorporated in the model, both to regulate the shift between groups and to define the total travelling population.

D LEONIDIS-PLESSAS, SC MATHEWSON
Department of Management Science
Imperial College of Science and Technology
Exhibition Road
London SW7 2BX
UK

A Rejoinder

IN RESPONSE to the note by D. Leonidis-Plessas and S. C. Mathewson, I would like to make the following clarifications, comments and possible concessions. Their last paragraph hints at the problem that has traditionally been faced when studying the pricing structure of the airline industry. That problem is that of elasticity, and many individuals represented by scheduled and charter airlines, regulatory bodies, consumer interest groups and independent academics have attempted to provide insight into the nature of elasticity. When the majority of the air travel market consisted of business traffic, the general conclusion was that the best estimate was represented by unit elasticity. Clearly the mix of passengers, especially in the North Atlantic market, has shifted to the tourist

and or vacation market. The elasticities representing price response for the business and tourist segments has been very difficult to ascertain, especially since the data or response to varying prices usually represents differing packages of services ranging from first class to standby on scheduled aircraft and from affinity to inclusive tours on charter aircraft. I attempted to approach the problem from the standpoint of estimating elasticity. But, because of the varying services, had to construct a more inclusive model that would also reflect differing product offerings. Elasticity was introduced where it was compatible with the more inclusive model (see footnote on p. 269 and the explanation of scheduled versus charter traffic on p. 274). Even here, however, a clear notion of elasticity is not evident because one really is observing the 'cross elasticity' between the charter and scheduled services in the estimates I have used.

Leonidis-Plessas and Mathewson are concerned about the scheduled group of passengers classified as 'other'. For the main part these fares (except the family plan which was only a small portion of the group) were for standby services. This group of people are not guaranteed a seat or a scheduled departure time. Thus, their inclusion in the scheduled group was somewhat arbitrary and is done only because these passengers do ultimately occupy a seat on the scheduled services. In spirit this group is not really interested in the maintenance of regular and dependable frequent departures offered on schedules but only a low price. They (as the charter passengers) will adapt their travel plans to the availability of the services. Meanwhile the typical passenger on scheduled services expects their reserved services to be provided in accordance with the published times and itineraries in the airline guide and identified on their ticket stubs.

In lieu of the above and the questions raised, further study is needed for the group defined as 'other.' Their price, while it was assumed to be of unit elasticity or similar to the higher priced scheduled traffic, is probably more elastic. Intuitively, however, it probably does not equal or exceed that of the overall charter market. Many of the military standbys are coming home on annual leave and are attempting to save dollars to spend elsewhere on the leave. Most of these individuals would probably pay the higher fares if the standby was not available. A major portion of these individuals are probably travelling off of 'peak' season and find that the standby provides them with adequate, although not totally predictable, service. Thus the elasticity for the 'other' group probably falls somewhere between the aggregate elasticity demonstrated by scheduled passengers and that demonstrated by the charter passengers.

As far as the differences between the figures taken from Fig. 5 and Table 1 we have to look again at the needs of the individuals in the 'other' group. If low military standby, student standby, and very special family package fares discontinued the potential passengers have these alternatives: (1) Pay a higher scheduled fare, (2) Make arrangements for travelling by charter, and (3) Not travel. The combination of the four curves indicate that 1 million of the 1·536 million would have selected either alternative (1) or (3). Figure 5 is derived from Figs. 2, 3 and 4, thus if you accept them you have to accept the result generated in Fig. 5. Because of the lack of sufficient hard data Figs. 3 and 4 were prepared using conservative estimates (see dashed line in Fig. 3); thus, one might speculate that the situation is even worse than demonstrated.

The above is a discussion that should shed some light on the nature of the model's development. I will concede that some of the arguments are not perfect and sometimes generate what appear (depending on your viewpoint) to be conflicting results. The model attempts to develop relationships and a basis for interpreting the data that relate to a very difficult problem.

WILLIAM G BROWNE

School of Business
Oregon State University
Corvallis, OR 97331
USA

Vol. 5
1977

Omega, Vol. 5, No. 1

A Management Science Approach to Planning Freedom—A Rejoinder

I APPRECIATE the opportunity of responding to Samuel Eilon's editorial in which he referred to my views on Planning and Freedom [2, 5]. I found little to argue in Eilon's statement although I regret that he mistakenly allowed himself to jump on the bandwagon by taking sides on the eternal polemic between those who prefer a planned system to an unplanned one, a free to a directed economy etc., and that he dwelled on the *differences* in philosophies among political and economic systems. That these differences exist, there is no question. However, my objective was to study the planning systems from which freedom (or lack of freedom) emerges, *qua* systems. As a system scientist manager and planner I would want to concentrate on the similarities among the planning systems rather than on their differences, in order to fashion a general systems theory which could apply to all planning systems regardless of their so-called political hue or tendency [6]. As a matter of fact, I find it imperative to avoid taking sides in the argument regarding which system is preferable—not that one could ever make that kind of rigorous determination anyway.

The assumptions–conditions–opportunities triad provides the key to an understanding of how freedom is determined in the context of the planning system, much like strategies and plans evolve in the context of strategic and corporate planning at the level of the firm [3].

Emphasis should be shifted away from a discussion of whether we, as clients of one system, like or dislike the results of the process taking place in another system and focused on the study of the process itself. In other words we need to observe the planning system from a level other than that at which the attitudes and philosophies (I called them 'assumptions') are thrashed out and are finally agreed upon. To plan, control and regulate the planning process itself, we must rise to the metasystem level as suggested by Beer [1] where the study of the planning system becomes a dispassionate study of how the forms of freedom result from the relationships among the components of the designing triad.

Planning "which culminates in the provision of choice" and the political–social system "which determines the exercise of choice", to use Eilon's apt description, are only parts or subsystems of the planning metasystem in which they are integrated [4]. The seemingly clear demarcation line separating provision and choice is as illusionary as that existing between amounts or degrees of freedom. Indeed we are dealing with 'wicked' and complex problems which can only be treated with a methodology which deals at the level of 'method' rather than of 'content', with 'process' rather than 'cause' and with 'structure' rather than 'substance'. It is only at that level that freedom becomes amenable to planning and that planning for freedom becomes meaningful.

REFERENCES

1. BEER S (1975) *Platform for Change*. Wiley, New York.
2. EILON S (1976) Planning and freedom. *Omega* **4**(5), 1–5.
3. MASON RO (1969) A dialectical approach to strategic planning. *Mgmt Sci.* **15**(8), B-403–B-414.
4. OZBEKHAN H (1969) Toward a general theory of planning. In *Perspectives of Planning*. (Ed. JANTSCH E), OECD, Paris.
5. VAN GIGCH JP (1976) Planning for freedom. pp. 47–158. *Mgmt Sci.* **22**(9), 949–961.
6. VAN GIGCH JP (1974) *Applied General Systems Theory*. Harper & Row, New York.

JOHN P VAN GIGCH

Operations and Systems Management
California State University
6000 J Street
Sacramento
CA 95819
USA

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

Judgmental Measures for Public Project Selection[1]

RONALD E TURNER
Queen's University, Kingston, Canada

(*Received January 1976; in revised form August 1976*)

An empirical study is described in which the feasibility of employing multidimensional scaling to obtain perceived measures of public projects is investigated. Data concerning eight possible muncipal projects were gathered from city officials and citizens of Kingston, Canada. Response problems are identified and the limitations for implementing the methods are discussed.

l. 5
77

INTRODUCTION

THE MUCH-publicized problems of the cities have focussed attention in recent years upon the managing of municipal organizations. Both elected and appointed officials have been urged to adopt management methods which have proven useful in the private sector [1]. Nowhere is this need more pressing than in the managing of public expenditures. Specifically, those budget decisions which determine which projects are undertaken by city administrations would seem to justify careful scrutiny to ensure that public funds are used prudently.

A concurrent trend which can be expected to affect the project-selection decisions in some city administrations is that of citizen participation. While the merit of receiving explicit inputs to the budgetary process from citizen groups can be debated pro and con, the fact of their increasing participation cannot be ignored [3]. One probable consequence of opening the 'black box' of project selection to public scrutiny is the need for more formalization of the procedures employed in making the selection decisions.

This emphasis on private-sector methods and citizen participation might seem to set the stage for the adoption of some of the formal techniques of preference measurement[2] which are used by marketing researchers. The applicability of

[1] The author gratefully acknowledges support for this study which was received from the Ministry of State for Urban Affairs, Ottawa. The cooperation of officials of the City of Kingston was also a valuable input.

[2] Preference measurement, as used in the first two sections of this paper, also includes the measurement of residents' perceptions of public projects. The distinction between the measurement of preference and perception is postponed for reasons of exposition until the measurement techniques are discussed in section three.

certain of these techniques is appraised in this study. Based upon psychometric methods for measuring peoples' perceptions and preferences, these techniques provide multidimensional measures of each project which appear well-suited for project-selection decisions. However, the types of judgments which are required to make the techniques operational raise questions of validity within the context of public projects, and suggest caution in any widespread attempt to apply them.

A municipal setting which provided an opportunity to appraise several of the modern marketing research techniques is described in the next section. This is followed by a brief description of each of the techniques and of the design of an exploratory experiment to compare them in practice. The results of the experiment are then discussed, and the study concludes with some recommendations for future applications.

SETTING

Vol. 5
1977

Kingston is a city of 60,000 residents situated at the edge of Lake Ontario about midway between Toronto and Montreal. By examining each issue of the local newspaper for the preceding 2 months, a preliminary list was prepared of projects in which local residents might express an interest. This list was further reduced in consultation with city budget officials by employing two criteria: the remaining projects should be of comparable dollar cost, and it should be feasible to initiate them individually within the next 3 years.

The following eight projects were chosen for inclusion in the study:

public library expansion;
airport improvements;
fire-fighting equipment;
storm sewer extensions;
Olympics (1976) sailing facilities;
garbage recycling plant;
parks and playgrounds;
improvement of a traffic artery.

Budget constraints precluded implementing all of these projects so this necessitated choosing among them on the basis of one or more relevant criteria. While costs and other technical criteria could not be ignored, it was felt that the principal choice criteria should be benefit-oriented. Objective measures of benefit are difficult to obtain in advance of implementation, however. Therefore, it seemed worthwhile to examine subjective measures such as perceived benefit from the viewpoint of potential users.

A traditional method for measuring perceived benefit might consist of asking residents to identify preferred projects, typically by presenting the projects one pair at a time. From such judgements, one could infer a ranking of the projects in order of their overall preference. This would be of limited use for making budget allocation decisions, however. City officials must weigh the benefits of one project against those of another project. Such tradeoffs require measures which indicate, not only which projects are preferred, but *how much* one project is preferred over another.

A second limitation is that the usual preference ordering is unidimensional—that is, the relative position of a project is based on one composite measure of preference for it. To appraise the benefit tradeoffs may require an understanding of *why* one project is preferred to another. If the underlying factors which determine a project's overall preference are to be evaluated, this means that the preference measures should be multidimensional.

A third measurement characteristic which is desirable is the ability to differentiate between residents based upon their evaluations of the projects. This can be important for elected officials who may want to know whether the electorate can be segmented into groups having similar preferences. A composite preference ranking precludes doing this except in a very crude manner.

Multidimensional measurement techniques which overcome the above limitations have been developed and applied in other fields. The literature of psychometrics and marketing research contain examples of the measurement of colours [10], cigarettes [7], and beer [5]. Although more complex entities such as business schools [4] and nations [12] have also been measured, it is not clear how suitable the measurement techniques are for public projects. One potential problem is the inherent diversity among the projects. Can a common reference frame be maintained by a respondent when comparing such diverse projects as parks and storm sewers? More generally, are the types of questions which must be asked to overcome the above limitations too complex or too time-consuming for the average city resident?

This study tries to provide an empirical answer to these implementation questions in a particular setting. But before describing the experimental design, it is necessary to provide additional description about the measurement techniques which are tested.

MEASUREMENT

To measure eight projects according to a unidimensional criterion obviously requires a vector of eight numbers. For present purposes, the scale on which these numbers are implicitly arrayed should have a higher metric than the ordinal; they should, at the least, possess the interval property, and preferably

the ratio property [11]. To measure the projects according to N criteria requires an N-by-eight matrix.

It is customary to represent the measured objects as occupying positions in a hypothetical N-dimensional space. By attributing Euclidian properties to this space, the position of each object can be specified by its N measures which play the role of a set of rectilinear coordinates.

Two types of space of this type have been hypothesized [4]. The above type containing only the measured objects has been termed a simple space. This represents the *perceptions* of the objects—specifically, the relative amounts which they possess of N attributes, each measured along one of the dimensions. The second type has been called a joint space since it contains both the objects and one or more persons whose judgments have measured the objects. The location representing such a person in a joint space represents his preferences for the objects. Conceptually, it is the combination of the measures which he considers is ideal, and for this reason it has been termed the ideal-point.

A technical explanation of the basis for locating the objects in a simple space is provided in the footnote references which follow. However, an intuitive explanation may be helpful at this point. Consider a respondent who is describing the relationships between a set of objects A, B, C, . . . , by stating his perceptions of the degree of pairwise similarity between them. For example, if he states that the pair AB is more similar than the pair AC, this would imply that A is 'closer' to B than it is to C. In a two-dimensional space, the three objects could occupy many locations which satisfy this single constraint. The number of possible locations would be narrowed somewhat if the respondent also stated his perceptions of the similarity relationship between pairs AB and BC, and that between AC and BC. The locations would be further constrained by introducing object D and judging the similarities between all of the newly-created pairs of objects. Indeed, as M the number of objects increases, a unique location for each object can usually be inferred from the $M(M-1)/2$ resulting pairwise relationships between them. Thus, each object can be simultaneously located on N scales having at least the interval property. An analogous rationale can be employed to justify the locations of objects in a joint space, except that the judgments from which the locations are inferred are based upon preference relationships rather than similarity relationships.

Figure 1 illustrates a two-dimensional simple space in which the locations of the eight projects were inferred from the similarities judgments of a group of Kingston residents. Each project can be characterized by a vector of two numbers which are its measures on two scales. The labelling of the scales in Fig. 1 was accomplished by an auxiliary procedure.[3] This consisted of correlating the scale measures of each project with a set of ratings of each project on a number

[3] The 'maximum-congruence method' [9] is used to interpret the semantic content of each dimension in relation to the project ratings on the specified criteria. The project measures, which are the coordinates of Fig. 1, were computed using the INDSCAL technique [2].

Vol. 5
1977

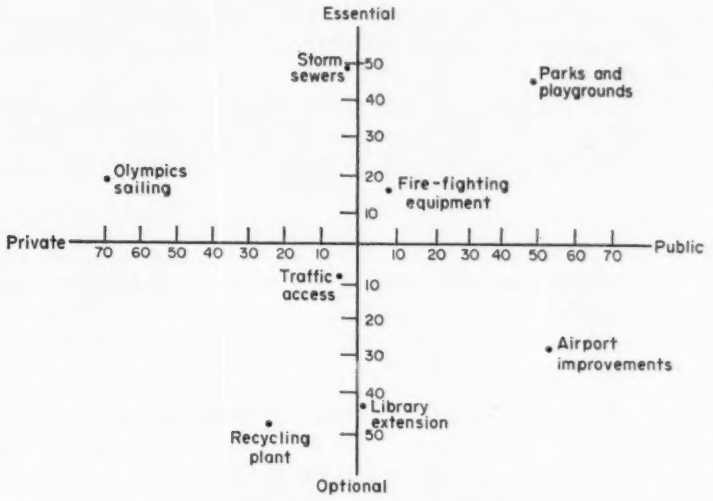


FIG. 1. *Relative amounts of implicit attributes of public projects.*

of criteria described in the next section. As a result, the perceived differences between the projects can be given a semantic interpretation. Thus, the traffic access and the storm sewers have comparable measures on the private–public dimension, but the storm sewers are seen as being the more essential.

The corresponding joint space would have a point added to Fig. 1 representing the amount of each dimension which a hypothetical ideal project would possess. Since the preferences of individuals usually differ, there would be a different ideal-point for each respondent. To whatever extent the ideal-points tended to form clusters, one might draw inferences about the existence of differentiable segments of preference. Elected officials might be interested in examining the demographic composition of the membership of such segments, for example.

Much more could be written about applications for the scaling models, as well as the details of how they work. However, the significant issue for present purposes is whether they make realistic demands on the residents who may be expected to use them as a vehicle for expressing their perceptions of the projects and their preferences for them. These judgments are described in the next section, after which an experiment for appraising their demands upon the citizens will be outlined.

RESIDENTS' JUDGMENTS

The measures of the eight projects as represented in two dimensions by Fig. 1 are derived entirely from certain judgments about the projects which were elicited from the residents. Different types of judgments are required for the

simple space represented in Fig. 1 than for the corresponding joint space. The two types of judgments are termed similarities judgments and dominance judgments, respectively.

The *similarities* judgments are intended to represent a resident's unstructured perceptions of the relationships between the projects, typically without any criteria or context being suggested. Two procedures for eliciting the similarities judgments are:

1. Prepare a deck of 28 plain cards, on each of which is the name of a different pair of the eight projects. Ask the resident to sort the deck in order of increasing similarity, such that the card at one end contains the names of the two least similar projects and the card at the other end contains the names of the two most similar projects.
2. List each pair of projects opposite a separate seven-point rating scale labelled 'low similarity' to 'high similarity'. Ask the resident to designate the amount of similarity between each pair of projects by placing a mark in the appropriate position on each scale.

Vol. 5
1977

Both of the above procedures are employed in this study to elicit similarities judgments from different residents. For ease of reference, the first procedure will be termed 'similarities by pairwise responses' (SP), and the second will be termed 'similarities by scaled responses' (SS).

The *dominance* judgments are somewhat more structured, in the sense that the respondents are asked to indicate their overall preferences for the projects. Two procedures for eliciting the dominance judgments are:

1. Ask the resident to identify the most preferred project. Consider in turn, each of the seven pairs of projects formed by comparing the remaining projects with the most preferred project, one at a time. For each such pair, ask the resident to divide 100 points between the two projects in proportion to his preferences for them. For example, if he preferred one project three times as much as the other, the appropriate division of the points would be 75:25.
2. Ask the resident to rank the eight projects by marking numerals representing the projects on an eight-point scale labelled 'most preferred' to 'least preferred'.

Both of these procedures are also used in this study. The first procedure will be termed 'preferences by pairwise responses' (PP), and the second will be termed 'preferences by scaled responses' (PS).

The above four procedures provide alternate methods for eliciting similarities judgments (to construct a simple-space map showing project perceptions), and

for eliciting dominance judgments (to construct a joint-space map showing project preferences). While each procedure provides the judgmental inputs which are necessary to measure each project's *location* along the dimensions, these judgments do not permit the *identities* of the dimensions to be inferred. To do this, a third set of judgments is required. This simply consists of structured ratings of each project according to each of several criteria, which seem to be relevant based upon prior exploratory questioning. The criteria are:

attractiveness;
benefit to the average citizen;
benefit to Kingston;
economic development;
personal interest;
security of Kingston citizens.

To obtain these ratings from residents responding to the PP procedure, the same procedure is repeated except substituting each of the above criteria for 'most preferred'. Similarly, for those residents responding to the PS procedure, the same procedure is repeated except using binary scales having labels corresponding to each of the above criteria. The ratings produced by these procedures are then correlated with the projects' coordinates on each dimension, in order to interpret the semantic content of each dimension.[3]

Thus, either the PP procedure or the PS procedure provides the additional ratings which are necessary to label the project maps. By combining either procedure with one of the SP or SS procedures, a complete set of judgments can be obtained. In this way, it is possible to elicit both similarities judgments and dominance judgments by alternate measurement procedures. The existence of the alternatives facilitates an appraisal of the demands which the various procedures make upon a resident in his role of respondent. An experiment for doing this is outlined next.

EXPERIMENTAL DESIGN

The overall purpose of this study is to appraise the applicability of several multidimensional scaling procedures for measuring residents' perceptions of, and preferences for, alternative public projects. A simplistic approach might be to choose one procedure and see whether a sample of residents can answer the associated questions. However, it seems desirable to learn more than whether responses can be elicited. To encourage participation, the residents must not experience undue difficulty in responding. To make the results meaningful, the responses must be of good quality, in the sense that they should not contain many inconsistencies, for example.

Since it is difficult to appraise such matters when a single procedure is used, it seemed appropriate to employ more than one procedure and make comparisons between them. Combining the two procedures for eliciting similarities judgments with the two procedures for eliciting dominance judgments, produced the following four experimental sets: SP/PP, SP/PS, SS/PP, and SS/PS.

Besides eliciting the judgments previously described for each procedure, the following three measures were employed to appraise the relative merits of the procedures:

1. time for the resident to complete a procedure;
2. the resident's perception of the difficulty of responding;
3. the quality of the responses.

The time measure was observed by the interviewer. The difficulty measure was a rating by the resident on a seven-point scale, obtained in a supplementary question at the end of the interview. The quality measure was the 'stress' [8] of the similarities judgments. This is essentially a goodness-of-fit statistic which is derived from the judgments as part of the measurement procedure. It measures the extent of intransitivities in a set of similarities judgments.

The experimental design consisted of randomly assigning a sample of residents to the aforementioned experimental sets. Since the research was exploratory, a small convenience sample of 34 residents of Kingston was employed. Twenty-seven of these were members of three service clubs, on whose behalf a donation of $100 was made. The remaining seven were officials of the City administration. The latter were arbitrarily assigned to the same experimental set. The demographic characteristics of the thirty-four residents are described in Table 1.

TABLE 1. MEMBERSHIP OF RESPONDENTS IN SELECTED DEMOGRAPHIC SEGMENTS (PERCENTAGES)

			Location		Occupation	
Age	under 30	28	North Kingston	46	Housewife	36
	30–45	36	South Kingston	4	Businessman	18
	over 45	36	West Kingston	32	Professional	11
		—	Centre Kingston	11	Secretary	11
		100	Suburban	7	Student	11
Sex	male	47		—	Tradesman	7
	female	53		100	Retired	3
		—				—
		100				100
Property	owner	57				
	non-owner	43				
		—				
		100				

RESULTS

Each of the four experimental sets produced an *N*-by-eight matrix of the coordinates of the projects in both a simple-space and a joint-space configuration, including the ideal-point coordinates. The results make it possible to characterize each project by a single number representing an individual's (or the whole group's) composite perception of the project. Alternatively, these one-dimensional perceptions can be expanded into as many as nine dimensions, depending upon the analyst's desire for parsimony versus completeness. Since sample results have already been presented in Fig. 1, these measures need be discussed no more at this point.

The mean completion times across respondents are shown in Table 2. They reveal that procedure SS took two-thirds as long to complete as procedure SP, and that procedure PS took two-thirds as long as procedure PP. Thus, the

TABLE 2. MEAN COMPLETION TIMES FOR FOUR FORMS OF QUESTIONS

Question form	Minutes to complete
PS	14.2
PP	22.3
SS	22.3
SP	33.0

scaling method provided a considerable time savings over the method of pairwise comparisons, both for similarities judgments and for preference judgments. Furthermore, one can speculate that the differences would become even greater for a larger number of projects, since the number of pairwise comparisons increases almost exponentially with the number of projects.

A second observation from Table 2 is that the dominance procedures (PS and PP) required approximately two-thirds of the completion time required by the similarities procedures (SS and SP). Preferences and similarities judgments tend to complement, rather than substitute for, each other. Therefore, these relative times do not argue for one method rather than the other, but they do provide guidelines for the relative magnitudes of portions of the questionnaire. For example, the number of specified criteria used in procedures PS and PP could have been increased to ten without making these more time-consuming than procedures SS and SP.

Regarding the residents' relative difficulty of responding, the percentage distribution of scores is shown in Table 3. These reveal that the similarity procedures were seen as more difficult than the dominance procedures. A weighted mean score suggests that the perceived difficulty of the scaling form was essentially the same as that of the pairwise procedure. Thus, dominance

TABLE 3. PERCENTAGE DISTRIBUTION OF SCORES FOR 'DIFFICULTY TO RESPOND' QUESTION

Score value		Question form PS	PP	SS	SP
Extremely easy :	1	0	0	0	6
Very easy :	2	6	14	0	0
Easy :	3	44	29	25	6
Just right :	4	31	36	25	18
Difficult :	5	19	21	38	70
Very difficult :	6	0	0	6	0
Extremely difficult :	7	0	0	6	0
Median score		3	4	5	5
Mean score*		3.63	3.64	4.43	4.46

* Assuming equal-interval property and weighting by score values.

TABLE 4. KRUSKAL'S STRESS FOR INDIVIDUAL PERCEPTUAL MAPPINGS OBTAINED FROM SIMILARITIES JUDGMENTS

Dimensionality	Mean stress	Random (0.05)	Random (0.10)	Random (0.25)
1	0.346	0.276	0.298	0.326
2	0.171	0.111	0.120	0.132
3	0.078	0.022	0.028	0.046
4	0.036	0.000	0.000	0.002
5	0.009	0.000	0.000	0.000

Vol. 5
1977

judgments were easier to make than similarity judgments, regardless whether the judgments were given as scale ratings or as pairwise orderings.

The quality of the results are implied by Table 4 which shows the actual mean *stress* values, along with those which would have been obtained using random data [6]. The differences are not statistically significant at acceptable levels of confidence. The inevitable conclusion is that the similarities responses are of rather low quality. This finding applies equally to the responses obtained from the SS procedure and those obtained from the SP procedure.

The low quality of the similarities responses is consistent with the respondent's perceptions of difficulty making these kinds of judgments. One can speculate as to the possible reasons for this. On the one hand, the eight projects are a diverse set. Without any guidelines as to which project attributes to consider when judging their attributes, this is an abstract task. On the other hand, most of the respondents were unaccustomed to conceptualizing, and probably found the similarity concept ambiguous and confusing.

CONCLUSIONS

Although this study was concerned with methods for obtaining measures to be used for project selection, the actual application of the measures is beyond

the scope of the study. However, it should be noted that all of the procedures produce measures having interval-scale metricity or better. Furthermore, the procedures all have the flexibility of providing, at the analyst's discretion, a single measure or a set of measures for each project. This potential to transform residents' judgments into measures seemed to justify an appraisal of the various measurement procedures.

Apparently, residents experience more difficulty making similarities judgments than they do making dominance judgments. Of the two types of output, perceptual mapping and preference mapping, the latter seems to hold the greater promise for public project selection. Furthermore, from the standpoint of efficiency, the scaling procedure for eliciting dominance judgments seemed to be less time-consuming than the method of pairwise comparisons.

Thus, the study suggests that preference mapping should be given serious consideration as a procedure for measuring residents' preferences for public projects. This provides a realistic point of departure for the further application of the procedure and the resulting measures in project-selection decisions.

REFERENCES

1. CAMPBELL AK (1970) What gets priority? *Nat. Civic Rev.* 466–72.
2. CARROLL J and CHANG J-J (1970) Analysis of individual differences in multidimensional scaling via an *N*-way generalization of Eckart–Young decomposition. *Psychometrika.*
3. FORT WORTH PLANNING DEPARTMENT (1974) *Citizen Participation and the Sector Planning Process.* Planning Department, Fort Worth, TX.
4. GREEN PE and TULL DS (1975) *Research for Marketing Decisions*, 3rd Ed. Prentice-Hall, Englewood Cliffs.
5. JOHNSON RM (1971) Market segmentation: a strategic management tool. *Jl Marketing Res.* 16.
6. KLAHR D (1969) A Monte Carlo investigation of the statistical significance of Kruskal's nonmetric scaling procedure. *Psychometrika* 319–330.
7. KLAHR D (1970) A study of consumers' cognitive structure for cigarette brands. *J. Bus.*
8. KRUSKAL JB (1964) Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika* 1–27.
9. MILLER JE, SHEPARD RN and CHANG JJ (1964) An analytical approach to the interpretation of multidimensional scaling solutions. *Am. Psychologist* **19,** 579–80.
10. SHEPARD RN (1962) The analysis of proximities: multidimensional scaling with an unknown distance function: II. *Psychometrika.*
11. STEVENS SS ed., (1962) Mathematics, measurement and psychophysics. *Handbook of Experimental Psychology.* John Wiley, New York.
12. WISH M, DEUTSCH M and BIENER L (1972) Differences in Perceived Similarity of Nations, in A. Kimball Romney, Roger N. Shepard, and Sara Beth Nerlove (eds.), *Multidimensional Scaling: Theory and Applications in the Behavioral Sciences*, Volume II, 289–313. Seminar Press, New York.
13. YOUNG F (1967) TORSCA Torgerson Multidimensional Scaling Program, University of Southern California.

ADDRESS FOR CORRESPONDENCE: *Professor RE Turner, School of Business, Queen's University, Kingston, Canada K7L 3N6.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

Insurance Exposure and Investment Risks: Management and Regulatory Implications

JOHN P MATTHEWS[1]
HOWARD E THOMPSON[2]
Graduate School of Business, University of Wisconsin

(*Received April 1976; in revised form July 1976*)

By exposing itself to greater risks on invested funds the company can also increase its expected returns. Because there are risks involved with both the insurance and investment operations, a major problem for an insurance company is to determine just how the two types of risks are to be balanced. In an earlier paper the authors dealt with the relationships among growth in premiums written, insurance loss and investment return. In this paper the questions are revisited from a policyholder and regulatory point of view, and suggestions are made concerning the reaction of the firm under various postures the management may take towards risk.

IN AN EARLIER paper, Thompson, Matthews and Li (TML) [2] developed a chance constrained programming model designed to analyze some of the rules of thumb which had been proposed by Kenney [1] for judging the strength of multi-line non-life insurance companies. In particular, the model was designed to deal analytically with the way in which risks with respect to insurance loss and investment return were related to growth in premiums written and the investment mix. From analysis of the model with hypothetical data, the rule of thumb that the ratio of premiums to surplus should be inversely related to the proportion of common stock in the investment portfolio was shown to be reasonable but not universally applicable.

This inverse relationship cannot be thought of as a rule to follow in prescribing policy or as an immutable measurement device to judge the financial strength of a company. Slow growth in premiums does not follow from a high common stock ratio; a high common stock ratio does not mean that steps should be made to reduce growth. Rather, both measures follow from dealing

[1] Assistant Professor of Business.
[2] Mary Rennebohm Professor of Business.

with the more fundamental risks in both insurance losses and investment returns.

Our previous analysis could be characterized as placing primary emphasis on the ability of the company to meet claims of policyholders. In this paper, we will shift the focus to the shareholder of the insurance company and deal with further aspects of the way in which risk relates to both premium growth and investment mix. We will assume that additional business can always be written provided we are willing to accept business with *increasing* marginal loss ratios. It is through this latter relationship that underwriting and investment are closely related and must therefore be closely coordinated for optimum results. Furthermore, we will vary the expected returns on the investments in order to recognize that, since both expected investment returns and the variance of returns are influenced by the state of the economy, both should affect investment mix and premium growth.

I. THE INSURANCE–INVESTMENT-MIX MODEL

A. The TML model [2]

The TML model consisted of

$$\max_{Y,X_1,X_2,X_3} E[A+B] \tag{1}$$

subject to $$Pr\{[S+(1-e)Y-\Sigma X_j \geq] 0\} \geq \beta_1; \tag{2}$$

$$Pr\,[(A+B) \geq 0] \geq \beta_2 \tag{3}$$

where $A = S + (1-\tau)(1-e-L)\,Y$, $B = \Sigma[(1-\tau)R_j + (1-\tau')U_j]X_j$ and $\Sigma = \sum_{j=1}^{3}$ and where Y is the premium written in dollars, X_j is the investment amount in security j, R_j is the realized rate of return on security j. U_j is the unrealized rate of return on security j, S is the surplus at the beginning of the period, τ,τ' are tax rates on realized and unrealized returns, β_j are probability specifications incorporating the willingness to accept risk, e is the expense loading and L is the loss ratio. Constraint (2) is the budget constraint, (3) is the liquidity–profitability constraint. R_j and U_j are assumed to be independent and normally distributed random variables.

B. The modified model

1. *Liquidity–profitability constraint modification.* In modifying the model, we will leave (2), the *budget* constraint, unchanged. Constraint (3) is basically oriented toward protecting the policyholder. Surplus is looked upon as an amount available to cover losses due to increased claims or poor investment results. An alternative (from the shareholder's point of view) is to consider surplus as the investor's earnings retained in the business. Although the policy-

Vol. 5
1977

holder may look at surplus as available to pay his claims and therefore is satisfied with a high probability that his claims will be paid (and that ending surplus will not be negative), the investor undoubtedly looks to a high probability that his maximum losses will not exceed some fixed proportion of the surplus.

The ideas of the previous paragraph lead to a new formulation of the constraint as

$$Pr\{(A + B) \geq (1 - C)S\} \geq \beta_2. \tag{4}$$

If $C = 1$, corresponding to the case where the investor would be willing to tolerate complete loss of the surplus with probability $1 - \beta_2$, then (3) and (4) are identical. If, on the other hand, $C = 0$, then the investor desires a probability of β_2 of having no loss, however small. Clearly, an insurance company decision on growth and investment mix must be made with *both* the policyholder and the shareholder in mind. Thus, both (3) and (4) are theoretically significant. However, when $C \leq 1$, (4) will always dominate (3) and we will therefore replace (3) with (4) in our subsequent analysis.

2. *Marginal loss ratios.* As previously indicated, it is likely that as more insurance premiums are written, the costs of the insurance will rise. We have chosen to let the cost increase be reflected in an increase in the marginal loss ratio. The simplest function which has an increasing marginal loss ratio is

$$L = L_1 + L_2 Y^2, \quad L_2 \geq 0. \tag{5}$$

The marginal loss ratio is then $\dfrac{dL}{dY} = 2L_2 Y$. For our analysis, we have assumed that $V(L_2) = 0$.[3] Therefore, using equation (5) the complete model may be written as

$$\max_{Y, X_1, X_2, X_3,} \quad E(A + B) \tag{6}$$

subject to (2) and

$$Pr[(A + B) \geq (1 - C)S] \geq \beta_2 \tag{7}$$

If we assume that L_1, R_j, U_j and e are normally and independently distributed, we can then convert equations (6), (2), and (7) to certainty equivalent form in a manner similar to those previously reported in [2]. The certainty equivalent has a non-linear objective function (because of Y^2) and is subject to non-linear constraints containing squares and cross products of X_j's, Y, and Y^2, as well as terms involving Y^4 and Y^6 and is considerably more complicated than the certainty equivalent to the model reported in section IA above. A statement of the certainty equivalent form and its derivation is available from the authors.

[3] Assuming that $V(L_2) = 0$ simplifies subsequent calculation but does not affect the essence of the model since $V(L) \neq 0$ incorporates variability of the loss ratio into the model.

II. SOME PSEUDO-EMPIRICAL EXPERIMENTS

The model developed in the previous section is sufficiently complicated so as to make the derivation of general conclusions by analytical methods all but impossible. Therefore, we used the model with hypothetical parameter values to derive conclusions on the relationships of investment yields, investment risks, and stockholder tolerance of loss to investment mix, insurance growth, and the marginal loss ratio of insurance written. We have termed this type of analysis pseudo-empirical, in that we will derive conclusions from experimental runs of an analytical model with hypothetical parameter values.

The economic conditions that an insurance company faces may be partly characterized by the relationship between the yields on the various investment possibilities. Bond yields may be high or low, and may in each instance be coupled with high or low expected stock yields. High expected bond yields and high expected stock yields would imply a situation similar to the conditions of late 1974 when the cost of funds was high and the stock market was reaching a low point. On the other hand, low bond yields and high stock yields or vice versa would be characteristic of periods of transition when the so-called yield curve is in a state of change; early 1975 was such a period. Each of the various combinations between bond and stock yields represents a particular situation that an insurance company may find itself facing.

Another economic factor which the insurance company faces is the degree of uncertainty with respect to yields. To translate uncertainty of future yields into operational form within the context of our model, we combine yield patterns with various levels of variance of yields.

The final factor which we investigate is the amount of surplus erosion that will be tolerated by the shareholder. The parameter C in (7) expresses the percentage of surplus erosion that will be tolerated with a small probability. Variations of this parameter will be combined with the expected yields and variances discussed in the previous two paragraphs in the experimental computer runs.

The variations in parameter values described in the preceding paragraphs are specifically detailed in Table 1, where the parameter values for the 54 experimental runs are listed. A few words describing the data are in order. In all runs, the risk acceptance parameters β_1 and β_2 are maintained constant at 0·975. The insurance parameters related to expenses and the *variance* of loss ratio are held constant in all runs. Furthermore, only the realized yield on bonds is varied; the unrealized yield and variances are held constant. Realized and unrealized yields on stocks are changed in the same proportion. There were two levels of expected stock yields, three levels of variance of stock returns, three levels of bond returns and three maximum loss factors employed which resulted in 54 different non-linear programs for analysis.

TABLE 1. PARAMETER VALUES FOR EXPERIMENTAL RUNS

Insurance	μ_e	0·25
Expense	σ_e	0·04
Parameters	σ_L	0·08
Investment	μ_{R_1}	0·07, 0.09, 0·11
Type 1	σ_{R_1}	0·00
Parameters	μ_{U_1}	0·00
(bonds)	σ_{U_1}	0·02
Investment	μ_{R_2}	0·03, 0·05
Type 2	σ_{R_2}	0·016, 0·020, 0·024
Parameters	μ_{U_2}	0·03, 0·05
(normal stocks)	σ_{U_2}	0·04, 0·05, 0·06.
Investment	μ_{R_3}	0·12, 0·20
Type 3	σ_{R_3}	0·008, 0·010, 0·012
Parameters	μ_{U_3}	0·072, 0·120
(growth stocks)	σ_{U_3}	0·12, 0·15, 0·18
Maximum loss factor	C	0·05, 0·10, 0·20

III. EXPERIMENTAL RESULTS

A. *Investment yields, investment risks, and investment portfolios*

Table 2 shows the investment mix for the 54 runs, with C set to 0·2.

By and large, the general results shown in the table are similar to those produced in [2] and in general what one would expect by financial theory. There are, however, some significant variations that merit some discussion. If we hold the yield on stocks at the low position and increase the yield on bonds —move horizontally across the table—the proportion of bonds (investment type 1) in the portfolio increases. This is true for every stock variance. Now, repeating the procedure with stock yield held at the high level, we find a somewhat different pattern. For the low variance case, the proportion in investment types 2 and 3 stays constant at about 40 and 60%, respectively. As we move to the medium and high variance cases, investment type 1 appears in the investment mix for the case of high bond yields. This result has a simple common sense explanation. In the case of low variance stocks, the yield on stocks is high enough when compared to bonds to cause bonds to be excluded from the investment mix. However, as the variance, or riskiness, of stocks increases, the lower expected yield on bonds becomes attractive because of the reduction in risk that bonds provide.

The patterns described in the previous paragraph are also evident as C is decreased but because the shareholder is increasingly less tolerant of a loss in

TABLE 2. % INVESTMENT MIX: SURPLUS EROSION TOLERANCE: $C = 0{\cdot}20$

Bond yield / Stock yield		Low			Low			High			Medium										
											Low			High			Low			High		
Investment type		1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3	1	2	3
Variance of stocks	Low	48·4	14·1	37·5	0	40·4	59·6	74·1	0	25·9	0	40·4	59·6	100	0	0	0	0	0	40·4	59·6
	Medium	62·8	11·3	25·9	0	58·1	41·9	82·3	0	17·7	0	58·1	41·9	100	0	0	0	20·3	38·0	38·0	41·7
	High	72·3	8·8	18·9	0	68·8	31·2	87·3	0	12·7	0	68·8	31·2	100	0	0	0	38·9	30·0	30·0	31·1

Vol. 5
1977

surplus, the previously described tendencies are more pronounced. Investment in bonds becomes, then, increasingly more predominant.

We can summarize the results on investment mix by stating that the relationships between expected yields on investment instruments, the variance of these yields, as well as the desire of the shareholders, will produce the investment mix of the firm.

B. Investment yields, investment risks, and insurance written

1. *Premium written as a method of dealing with risk.* Investment mix changes may not be sufficient to deal with the inherent risks in both insurance and investment opportunities. We wish to point out here that the level of and profit on insurance written as well as the variance of insurance losses affect the method of dealing with risk. These factors combine with investment yields and investment risks to determine the premium volume.

In [2], we previously illustrated the way in which the variance of insurance losses affects premiums written. Table 3 illustrates how insurance profit affects premium volume. As a preface to discussing the table, it is worthwhile noting that the relationships between the risk variables (C, β, and variance) and insurance written is not clearcut nor easily explainable, as was the case of investment mix. What is more, explanation of premium volume results will require careful consideration of the ability of investment mix decisions to deal with risk.

TABLE 3. PREMIUM–SURPLUS RATIO AND 100 × MARGINAL AFTER-TAX PROFITS ON INSURANCE ÷ PREMIUM WRITTEN. SURPLUS EROSION TOLERANCE: $C = 0{\cdot}20$

Bond yield		Low		Medium		High	
Stock yield		Low	High	Low	High	Low	High
Variance	Low	3·83	4·52	4·47	4·52	5·20	4·51
		−0·45%	−1·64%	−1,55%	−1·64%	−3·02%	−1·64%
of	Medium	3·89	4·39	4·52	4·39	5·20	4·54
		−0·55%	−1·41%	−1·64%	−1·41%	−3·02%	−1·68%
stocks	High	3·95	4·16	4·55	4·16	5·20	4·59
		−0·64%	−0·10%	−1·71%	−0·10%	−3·02%	−1·79%

Table 3 displays in each cell the premium–surplus ratio and 100 X marginal after-tax profits on insurance divided by premiums written for each of the 54 computer runs. Consider Table 3 for the case of low yielding stocks. When bond yields are high, the premium-to-surplus ratio is 5.20. As the variance of stocks increases, the premium surplus ratio does not change since the investment portfolio is one of complete investment in bonds. Thus, the variance of stocks has no influence on risks that the company faces and therefore on premiums written. Turning now to the marginal after-tax profit on insurance

(the second entry in the cell), we note that there is a marginal loss on insurance. The company is acting in the best interests of the shareholder by taking unprofitable insurance at the margin and investing the funds in the high yielding bonds.

Next, look at the case of low yielding bonds and low stock yields. In this situation, we see that the company is in a mixture of bonds and stocks. As the variance of stocks increases, the investment mix contains a higher proportion of bonds. This has the dual effect of reducing the expected yield from the investment portfolio but of also reducing the variance of profits from the investment portfolio. Table 3 shows that as the variance increases the premium written *increases*. This is because it is in the best interests of the shareholder to cause a switch to bonds, thus decreasing variance of profits, but at the same time increasing investment profits by supplying more funds for investment. The loss on the marginal premium is offset by the increase in investment yield that it will provide. And in addition, this is accomplished without causing any increase in the probability of having a 20% or more surplus drain. This result follows because the difference in yields between bonds and stock yields is not substantial, while the difference in variance is.

Since on all computer runs the constraints were active, if variances of investment returns are *increased* from one run to the next, the optimal solution from the previous run would no longer satisfy equation (7). The *previously* optimal values of the variables would continue to yield the same *expected* profit but as the investment return variances increase, the probability of loss would exceed $1-\beta_2$. Thus, in a sense, it is desirable to obtain a solution which satisfies the probability-of-loss (less than CS) constraint with higher investment return variances and keep the reduction in profits to a minimum. There are several possibilities to consider.

If the investment in growth stocks is reduced, variance is reduced. Depending on the reduction in variance, it may be possible to substitute bonds for growth stocks on a dollar-for-dollar basis and still reduce profits variance while doing so. If this dollar-for-dollar substitution is employed, premium volume may remain unchanged because the total amount of investment will remain constant.

Consider now the aspects of a situation where the additional investment in bonds can be a *multiple* of the reduction in stocks and still the variance in profits would be reduced. If it is possible to increase the investment in bonds *more* than the reduction of investment in growth stocks, then additional funds must be acquired. These funds would come from the insurance side and additional premiums would be acquired to provide the funds for the additional investment in bonds, thus increasing total investment.

This phenomenon, increased investment variance, may lead to a drastic investment mix change *and* an increase in premiums as observed in Table 3. As variance is increased from low to high, bonds are substituted for growth stocks. Premiums increase concomitantly, although the increase in bonds and premiums

does not entirely offset the reduction in expected returns from the reduction in growth stock investment.

Examination of the column denoting low bond yield and high stock yield in Table 3 shows a declining premium–surplus ratio as the stock yield variance increases. Consider first the low stock yield variance case. Because of the high yield on stocks and the low yield on bonds, no bonds are included in the investment mix, as shown in Table 2. When the stock variance increases, an adjustment in investment portfolio or insurance written must take place to maintain the probability of violating constraint (7) at 2·5%. One way to make the adjustment is to include more bonds in the investment portfolio. When this is done, both the expected return from the investment portfolio and the variance of that return decrease. In this instance, the reduction in expected return more than offsets the reduction in variance, thus *increasing* the probability of violating constraint (7). Another way to make the adjustment is to reduce the insurance written. This will have the effect of providing fewer funds for investment, thereby reducing the variance and expected return from investment. Notice that if we cut back premiums and investment sufficiently far there would be little chance of violating constraint (7). There are, however, other benefits from this strategy. First, we will stay in the higher yielding securities. Second, a cutback in premium volume will mean a decrease in the marginal loss ratio, thereby adding to profit. The third way of dealing with the increase in variance would be to change the mix between the two types of stocks at the same time that insurance written was decreased. This latter approach was the way chosen by the programming model.

2. *Capital structure, growth, and risk.* Reserves created by insurance premiums can be looked at as analogous to debt or borrowed funds. The surplus is, of course, the equity capital. If one looks at an insurance company as a financial intermediary, the analysis we have performed can be viewed as relating to the capital structure of the intermediary. Since the policyholders' surplus is fixed in our analysis, changes in premium–surplus ratios denote changes in the optimal capital structure. What is more, this optimal capital structure is related to the investment opportunities available to the insurance company. It is further related, and most importantly, to the willingness of the shareholder to tolerate a surplus erosion.

Table 3 shows marginal insurance losses in all cases. The willingness of the shareholders to tolerate the possibility of surplus erosion means that they are willing to expand the debt and pay a positive cost of that marginal debt in the hope of realizing large returns on their investment portfolio. When C is reduced significantly, there will be no marginal losses on insurance. This corresponds to a *negative* marginal debt cost. Interpreted in another way, it indicates that the risks of this financial intermediary are such that it is only worthwhile to add an additional increment of borrowed funds if the company is paid to do so.

IV. SOME POLICY IMPLICATIONS

1. *Management of growth and investment.* The analysis of the model makes it clear that the values of C and β_1, β_2 have important effects on the expected returns to the shareholders of the insurance company as well as the riskiness of those returns. It is well known that specification of the β's implies a degree of risk aversion on the part of the decision maker. Very large values of β_1 and β_2 would correspond to a high desire to avoid risk, whereas small values of β_1 and β_2 would imply a willingness to accept more risk. It also follows that large values of β_1 and β_2 would be associated with small expected returns. Thus, if the shareholders wanted greater expected returns, they would have them by reducing β_1 and β_2, but in the process would accept more risk.

The value of C is another risk acceptance parameter. The larger the value of C, the more willing the shareholder would be to tolerate some erosion of the surplus. Clearly, this mean that they are more willing to accept a higher probability of a substantial loss. In general, as C is increased, the expected returns will increase, but the probability of negative returns will also increase.

2. *Regulation of insurance companies.* We indicated in section IIIB2 that insurance premium reserves could be viewed as comparable to debt. Our analysis had implications for the capital structure of the intermediary. Both the inherent risks in insurance and investment possibilities and the willingness of the shareholders of the company to tolerate potential surplus drain affected the capital structure. In terms of the model presented in this paper and the analysis in [2], we can conclude that as C increases, shareholders can increase expected gains but will do so only by tolerating a higher *potential* surplus drain. Furthermore, by tolerating a smaller β_2—that is, a large probability of exceeding the surplus drain tolerance—they can also increase expected gains. In both of these cases, the greater expected gains will come from increasing insurance written and therefore relaxing underwriting standards. Thus, the debt–equity ratio will increase.

The results of this analysis are not surprising when one considers the role of equity capital in the growth of *any* business firm. Because financial institutions perceive risks of losing funds, they will not make loans to businesses that have large debt–equity ratios. The growth of small business if then limited by their ability to secure added equity capital. The analogy between any business firm and an insurance company breaks down at this point. Business firms that borrow from financial institutions borrow from a sophisticated investor that can judge the safety of his loan. The lender to the insurance company is the policyholder, who has little interest in the leverage of the company and almost no ability to judge whether the risk of his receiving payment on a claim is substantial or not. Whereas the capital structure of the business firm is effectively controlled by the lending market, the capital structure and growth of an insur-

Vol. 5
1977

ance company cannot be so controlled. The insurance commissions are charged with the task.

Insurance commissions, being concerned with the financial ability of insurance companies to protect the policyholder, have devoted a great deal of attention to the establishment of rates. This frequently takes the form of making rates so that insurance profits will be maintained at some pre-determined level for an industry group. That is, a floor is placed on rates.

The rate-making process seems to be a sensible one, but cannot be regarded as necessarily one which will protect the policyholder. To expand on this point, assume that we have an insurance company that has a high value of C and a low value of β_1 and β_2. Because the shareholders are willing to tolerate substantial risks, the insurance company will optimally take on a high premium-to-surplus ratio; in other words, will take on more debt. Marginal losses on insurance will be likely. These marginal losses will cause the insurance profits to fall. If rates are raised so as to assure a reasonable profit on insurance operations to 'provide financial stability' for the policyholder, we may find that the policyholder pays more for the same financial stability he had before the increase. Presumably, if the company does not change C or β_1 and β_2, they will tolerate the same probability of surplus erosion but will do so at a greater volume of insurance (because the marginal loss ratio is lower for every value of premium volume), and at a greater debt–equity ratio. Little in the direction of financial stability has been provided by the rate regulation. The shareholder would be expected to reap large benefits in this situation with no increase in risk of financial loss.

V. SUMMARY AND CONCLUSIONS

In this paper, we have used a chance constrained programming model to study the way in which premium growth and investment mix are related to both insurance and investment risks as well as the costs of insurance—as expressed by the marginal loss ratio—and the structure of expected returns on investments. Varying investment mix and premium growth are ways of responding optimally to risk.

The premium reserves which result from the premiums written by the company constitute an analogue to debt in industrial firms. Therefore, the insurance and investment risks, the marginal costs and yields, as well as the tolerance of the shareholders for surplus erosion, can be thought of as determining the optimal capital structure in an insurance company. Growth and capital structure are two sides to the same coin.

In the absence of any explicit regulation, growth of the insurance company could be maintained at almost any rate desired, so long as the shareholders are willing to tolerate the associated risks. This results because of policyholders'

inability to judge the riskiness of their investment in an insurance policy. This is in sharp contrast to growth in an industrial firm which will be limited by the suppliers of capital who will accurately perceive their risks and refuse to lend to the company.

Regulation is needed, then, because of the difference between lenders to industrial firms, which are primarily sophisticated financial institutions, and lenders to insurance companies, which are primarily unsophisticated individuals. But rate regulation, designed to maintain reasonable profits for insurance companies, will most likely have little effect on the financial stability of the insurance industry. Only by effecting the willingness of shareholders in insurance companies to tolerate risk can the financial stability be affected.

REFERENCES

1. KENNEY R (1967) *Fundamentals of Fire and Casualty Insurance Strength.* The Kenney Insurance Studies, Dedham, MA.
2. THOMPSON HE, MATTHEWS JP and LI BCL (1974) Insurance exposure and investment risks: An analysis using chance constrained programming. *Ops Res.* **22**, 991–1007.

ADDRESS FOR CORRESPONDENCE: *Mr John P. Matthews, University of Wisconsin, School of Business, 1155, Observatory Drive, Madison, Wisconsin 53706 USA.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

Interdependence and Innovation[1]

A GERSTENFELD

Worcester Polytechnic Institute, Massachusetts

(*Received June 1976; in revised form July 1976*)

This article analyzes the relationships beyond the firm for their effects upon innovation. It is argued that as technology becomes increasingly complex close interfirm cooperation will be needed. In this study, twenty-two innovations were examined particularly focusing on the interaction between the firm and the environment. It was found that there was a significant correlation between innovation success and the firms' use of other firms. The role of government, universities, and outside inventors on innovation within the firm is also discussed.

1. 5
77

I. INTRODUCTION

THE IMPORTANCE of the organization's dependency on external relationships is receiving increased attention by authors. Similarly the subject of innovation and the management of R&D has been the focus of many studies. The purpose of this paper is to attempt to combine those two thrusts and consider the subject of innovation by focusing on the firm's external linkages.

There are many examples in which the individual firm is no longer able to have the necessary technological breadth essential for the development of current complex products and processes. In the US a close working relationship between Kodak and Sylvania was necessary for the development of a highly reliable flash cube for use with the 'instamatic' camera. Similarly a close organizational interdependence exists today between the Zenith Radio Corporation and the Corning Glass Works for the development of a widely discussed new color TV picture tube. This R&D project has involved an investment in excess of $15 million to date.

I mean in no way to underestimate the importance of the firm's internal organization and its effects upon innovation. Clearly, firm size, structure, R&D intensiveness, source of ideas, degree of internal integration, as well as other factors, are vitally important in affecting success or failure for R&D. What I have set out to do in this paper is to add another dimension which I feel has not received the attention that it should, namely the effects of external linkages upon innovation within the firm.

[1] Presented at the combined meeting of the Operations Research Society and the Institute of Management Science, April 1976, in Philadelphia, PA.

In doing so, I shall examine 22 innovations and trace the relationship beyond the firm to outside entities. The particular external linkages that I shall consider are:

- the firm's relationship with a university or consulting organization on a particular R&D project;
- the firm's relationship with a government agency or laboratory on a particular R&D project;
- the firm's relationship with other firms on a particular R&D project;
- the firm's relationship with an outside independent inventor on a particular R&D project.

The use of these external sources will be related to successful and unsuccessful projects to enable the reader to focus on the possible value of particular inter-organizational relationships for innovation.

II. PREVIOUS RESEARCH

Many previous studies have emphasized the importance of understanding communication barriers within the firm [9]. For example, I reported earlier on project failures due to poor communications between marketing and R&D [4, 5]. Similarly, Mansfield states, ". . . there is considerable evidence that the probability of commercialization (given technical completion) is directly related to the degree to which R&D and marketing are integrated." [8] It has perhaps now become time to examine the importance of integration beyond the firm.

Rothwell *et al.* [11] reviewed the SAPPHO data and among other findings showed the aggregate results for 22 pairs of innovations in the chemical processes industry and 21 in scientific instruments. On the subject of inter-organizational linkages the article shows that there was a significant relationship between the successful innovations and those that used outside technology during production. Similarly, Rothwell *et al.* showed the positive relationship between successful firms and those that have better external communication.

In another study, Rothwell and Robertson analyzed nine cases and focused on the contribution of poor communication to innovative failure [12]. Most of the cases cited showed poor use of outside sources as well as poor internal communication. One case involved an independent inventor as the inter-organizational link and the failure is similar to an example cited later in this paper.

A recent paper by Pavitt and Walker summed up three studies of innovations undertaken in the UK which focused on the firms' use of outside sources [10]. In two of the three studies it was found that inter-industry relationships were predominant.

Von Hippel found in a recent MIT study that the organizational link between user and supplier indicated that the innovation often came from the user rather than the supplier [19]. That study consisted of an examination of industrial

good innovations from four important types or families of scientific instruments. Scientific instruments were selected to be studied primarily because previous research on the innovation process had ascertained that innovation in response to user need was prominent in scientific instruments [17]. In the Von Hippel study, it was found that the instrument manufacturer utilizes information from the user, and that commercially successful innovations are often derived from prototypes built by the users. It is believed that scientific instrumentation is not the only industry for which this user-dominated pattern is typical. The importance of the inter-firm dependency is stressed.

Johnston and Gibbons point out that the most significant and often cited finding on relationships beyond the firm has been the importance of communication through personal contacts [6]. In my earlier text, I argue for porous organizational boundaries [2]. The organization must receive information from the environment and in turn the organization has effects upon the environment.

Mansfield points out that an enormous amount has been written concerning the management of industrial research and development [8]. He emphasizes that the rate of technological change could be increased materially, without any substantial increase in R&D expenditures, if intrafirm barriers to the utilization of R&D results could be reduced. I would support that argument but, indeed, would like to take it a step further and propose that the reduction of interfirm barriers can likewise lead to more effective results from R&D.

It has been pointed out by writers that long-term R&D commitments can be taken on more easily by large firms than by small [18]. The small firms in the future will find it increasingly difficult to afford the required R&D expenditure for innovation. I would argue that inflationary forces as well as increased technological complexity, will encourage interfirm R&D and that the successful innovators will be those who adopt such strategies.

III. DATA COLLECTION

The data was gathered by the author while teaching in West Germany and partially sponsored under an NSF grant to the MIT Center for Policy Alternatives. I limited my sample to the three most R&D intensive industries in West Germany; namely, automotives, chemicals, and electronics. The total number of companies interviewed was 11, although the total number of interviews was 55 or an average of 5 persons per firm. The interviews were with R&D managers, engineers, and scientists.

While I used a carefully constructed and tested interview guide in the field research, the range of questions usually discussed during the interviews was more encompassing than that included in the formal interview guide. My work was part of a large Five Country Study being conducted by MIT, but for the purposes of this paper, I would like to focus only on the subject of the relation-

ship between specific innovations and the firms' use of external sources. The distribution of firms I investigated consisted of the two largest chemical companies in West Germany, five automotive companies, and four electronics companies.

The total number of innovations discussed during the interviews were 22, or 2 per firm and were generally more developmental than research oriented. The critical incident approach was used, i.e. each firm was asked to describe a successful and an unsuccessful innovation. For any number of reasons, respondents might consciously or unconsciously bias their results when asked their opinions concerning innovation and a firm's external linkages. To reduce this bias, respondents were not asked for their opinions, but rather were asked to recount what had happened in a specific incident. Innovation was defined as a technological change that is new for the company.

For the definition of successful, the innovation was to have been complete, the product or process on the market, or in use, with every expectation of continuing sales [13]. The unsuccessful innovation was to have been a project in which a minimum of one man year was expended and the project is now stopped with no expectation of continuing. Each of the innovations was to have taken place recently with no project starting more than five years ago. The innovations are not necessarily representative. However, the critical incident approach does permit one to examine certain relationships.

Vol. 5
1977

Although many questions were asked regarding the relationship between the firm and the external environment for a particular innovation, the key questions were as follows:

- For this particular innovation, was outside assistance sought from any of the following:
  a. Universities and consulting organizations?
  b. Government organizations (agencies and laboratories)?
  c. Other firms?
- Did the idea for the project come from an outside inventor?

IV. RESULTS AND DISCUSSION

Universities and consulting organizations

Only one of the 22 innovations used assistance from a university. In my discussions with the firms, there seemed to be general agreement that there is little direct linkage between the firm and the university. The one case of direct technical assistance from a university to a firm involved the design of a turbo motor modification requiring testing in a large wind tunnel. For this work, the university facilities and faculty assistance was used. This use of university facilities by firms in West Germany appears to be generally infrequent, although it is not unusual for a company to invite a university professor to speak to a technical group regarding some general principles.

None of the innovations investigated used assistance from technical consultants. In my discussions, it appeared that the companies felt that they had

more expertise inhouse, and, therefore, seldom used consultants for technical problems. (It was, however, pointed out that the use of consultants for management issues is common.) It must also be remembered that the companies sampled were large and that for smaller companies, it is possible that different findings might have resulted.

Government organizations (agencies and laboratories)

None of the 11 successful innovations used the help of a government agency or laboratory, and none of the 11 unsuccessful innovations used government assistance. When I explored this more deeply, the explanations seemed to be that the large West German firms have good research facilities and there is seldom need to go to the government for assistance on the development of new products or processes. The most probable implication of this is that the more basic work being performed by the government was of little applied use for the firms' new product and process developments.

Other firms

It was found that outside assistance from other firms was used often and was associated with successful projects in nine out of eleven successful innovations (see Table 1). Needless to say, this may not be cause and effect, but the data

TABLE 1. ASSISTANCE OF OTHER FIRMS IN INNOVATION

	Project success	Project failure	
Use of other firms	9	1	10
No use of other firms	2	10	12
	11	11	22

The Fisher exact probability test shows $P < 0{\cdot}005$.

seems to show a high use of other firms which is statistically significant. These findings are consistent with Allen's findings showing the importance of outside firms to innovation [1]. The use of other firms by the innovating firm was generally a substantial part of the total research effort.

The independent inventor

The role of the independent outside inventor as an organizational link affecting innovation was also investigated. It is known that today's research efforts require teams, and it has been argued that most innovation comes from the use of cross disciplinary efforts [15]. Of the 22 projects I examined, only two came from outside inventors. Of the two, both were unsuccessful innovations.

During the presentation of these findings at a leading German university (University of Saarlandes), it was suggested that perhaps this low use of outside

inventions is due to the fact that the companies sampled were large companies. The argument was that large companies can afford their own inventors and research groups so one might expect to find a low frequency of outside inventors. That is, of course, possible, and one can only conjecture what the results would be with smaller companies. My belief is that the use of independent inventors is very limited although perhaps not quite as limited in small companies as in larger ones. It has been found in other studies that 80–90% of all inventions are made by employed inventors [14].

Mansfield states, "... independent inventors continue to produce a significant share of the important inventions, although their relative importance seems to have declined." [7] I found that the decline is even more significant than suggested, although none of the companies I interviewed were willing to close the doors to the possibility of the use of outside inventions.

As another issue regarding the use of outside inventors, it is important to consider that the invention is but a small part of the innovative process [3]. As such, it represents a small portion of the cost of innovation. Since innovation is defined as a given technological change, innovation can exist with or without invention. Schmookler points out that "Society generally uses far more of its resources to disseminate technology than to advance it. This is plain from the far greater manpower, including that of students, devoted to formal and informal technological education than to discovery and invention." [16] We can distinguish four kinds of technology-producing activities: research, inventive activity, development, and marketing. It is clear that the inventive activity may or may not be present in innovation, but the major costs are generally in development and marketing. The firms may, therefore, hesitate to invest the larger costs on a technological change from an outside invention in which they have limited confidence. This is rightly so since the 'in-house' invention is better understood by the personnel. Perhaps equally important is the strong commitment that is concomitant with one's own invention. Funds can then be invested in the more expensive portions of the innovative process.

The well established concep. of internal commitment also adds to the difficulty faced by the firms when trying to use outside inventions. It is too often that this commitment is lacking when the invention is from the outside. In my investigation of the cases in which an outside invention was used and the project was unsuccessful, I had further evidence of the 'not-invented-here' syndrome which I am afraid is still often used as the excuse for failure.

V. SUMMARY AND CONCLUSIONS

This study focuses on the use of interorganizational links and innovation. It was found that only a very small number of the projects studied used the direct assistance of universities or consulting organizations. Furthermore, none of the projects investigated used the assistance of government laboratories.

Vol. 5
1977

However, it was found that the use by the innovating organization of assistance from other firms was frequent and correlated with success.

The strong inter-organizational link, namely interfirm, was associated with a large portion of the research effort. The decision to seek assistance from another firm was generally based on skills which the second firm possessed which were not within the expertise of the innovating organization. One could argue that although these interfirm relationships may have been important for some time, they become even more important today as the individual firm is unable to maintain the breadth essential for current technologies.

The relationship between the independent outside inventor and the innovating organization was examined. To no surprise it was found that the use of outside inventors is limited, and it was surmised that, as technology becomes more complex, this link may be expected to become even further diminished.

It is believed that these findings coupled with other investigations, are starting to shed some light on the relationships, and perhaps place more emphasis on the importance of inter-organizational linkages and innovation. Like many studies, several important questions remain unanswered and will hopefully be the subject of future investigations.

l. 5
77

REFERENCES

1. Allen TJ and Cohen SI (1969) Information flow in research and development laboratories. *Admin. Sci. Q.* **14**, 12–19.
2. Gerstenfeld A (1970) *Effective Management of Research and Development*. Addison-Wesley, Reading, MA, USA.
3. Gerstenfeld A (1976) *Innovation: A Study of Technological Policy*. University Press of America, Washington, DC, USA.
4. Gerstenfeld A (1976) A study of successful projects, unsuccessful projects, and projects in process in West Germany. *IEEE Trans. Engng. Mgmt.* **EM23**.(3), 116–123.
5. Gerstenfeld A, Turk CD, Farrow R and Spicer R (1969) Marketing and R&D. *Res. Mgmt* **12**(6), 409–412.
6. Johnston R and Gibbons M (1975) Characteristics of information usage in technological innovation. *IEEE Trans. Engng Mgmt* **EM22**(1), 27–34.
7. Mansfield E (1971) *Technological Change*. Norton, New York.
8. Mansfield E and Wagner S (1975) Organizational and strategic factors associated with probabilities of success in industrial research and development. *J. Bus.* **48**(2), 179–198.
9. Meyers S and Marquis D (1969) Successful commercial innovations. *National Science Foundation*, NSF 69–71.
10. Pavitt K and Walker W (1976) Government policies towards industrial innovation: a review. *Res. Policy* **5**(1), 15–93.
11. Rothwell R, Freeman C, Horlsey A, Jervis VTP, Robertson AB and Townsend J (1974) SAPPHO Updated—project SAPPHO Phase II, *Research Policy* **3**, 258–291.
12. Rothwell R and Robertson AB (1975) The contribution of poor communication to innovative failure. Workshop at Polytechnic of Central London.
13. Rubenstein A, Chakrabarti A and O'Keefe R (1974) Final Technical Report on Field Studies of the Technological Innovation Process. Northwestern Univ., USA.
14. Schade H (1972) Employees' inventions—law and practice in the Federal Republic of Germany. *Ind. Property*, 249–254.
15. Scherer FM (1967) Market structure and the employment of scientists and engineers. *Am. Econ. Rev.* **57**, 529.

16. SCHMOOKLER J (1966) *Invention and Economic Growth.* Harvard University Press, Cambridge, MA.
17. UTTERBACH J (1971) The process of innovation: a study of the origin and development of ideas for new scientific instruments. *IEEE Trans. Engng Mgmt.* 124–131.
18. VERNON R (1971) *Sovereignty at Bay.* Basic Books, New York.
19. VON HIPPEL E (1975) The dominant role of users in the scientific instrument innovation process. *MIT Working Paper*, 764–75.

ADDRESS FOR CORRESPONDENCE: *Professor A Gerstenfeld, Head of the Department of Management, Worcester Polytechnic Institute, Worcester, MA* 01609, *USA.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

Quadratic Programming Applications[1]

BRUCE A McCARL

HERBERT MOSKOWITZ
Purdue University

HARTLEY FURTAN
University of Saskatchewan

(Received February 1976; in revised form June 1976)

Although quadratic programming (QP) is often studied as a methodology by decision scientists, little emphasis seems to be given to the type of decision science problems that can be so modeled. This paper reviews and extends some of the applied and methodological areas where QP is applicable, discussing and illustrating the characteristics and aspects of the accompanying solutions. Some of the methodological areas amenable include regression analysis, decision analysis, and quadratic approximations to generally complex functions. In the functional management areas, QP is applicable to problems in economics, such as demand–supply response and enterprise selection. In finance, it is used in portfolio analysis; in agriculture, in crop selection. Both the methodological and functional applications of QP are reviewed. Potential new applications of QP are suggested.

QUADRATIC programming problems (QPPs) have long been of theoretical interest to management and decision scientists. While many management and decision science educational programs cover quadratic programming (QP) theory, little emphasis is given to possible empirical applications. Moreover, a unified presentation of existing quadratic formulations is not found in any one body of literature, either journal or text. The purpose of this paper is to review existing applications of QP to various areas in the management and decision sciences, to point out, where applicable, the added insights and capabilities obtainable by such formulations, and to suggest potential new formulation areas. While the treatment of any one application in this paper is not by any means complete, the authors feel that the discussion, examples, and bibliography should aid educational, research, and application endeavors.

[1] The research reported in this paper was completed as Journal Paper 5803 under Project 1788 of the Purdue Agricultural Experiment Station.

Quadratic programming theory

The theory of QP was initially developed for the problem of maximizing (or minimizing) an objective function consisting of a linear term plus a quadratic term, subject to (s.t.) some linear constraints. Symbolically, we have

$$\text{Max.} \quad Z = C'X + X'\phi X \tag{1}$$

$$\text{s.t.} \quad AX \leq 0 \tag{2}$$

$$X \geq 0. \tag{3}$$

The quadratic term $X'\phi X$ is most frequently a symmetric negative semidefinite (concave) matrix. Hence, $X'\phi X \leq 0$ and reflects the interactions among the decision variables. The degree of interaction is manifested by the coefficients in the ϕ matrix.

Solution procedures for QPPs generally rely upon the application of the Kuhn–Tucker theory. Many algorithms have been developed for solving QPPs. Boot [1] and Wagner [23] review some of these algorithms, and provide a list of references. Associated with QP solutions are activity levels, marginal activity costs, resource disposal levels and resource shadow prices. These items of information are similarly interpretable as in the linear programming (LP) case.

Vol. 5
1977

APPLICATIONS

The remainder of this paper deals in detail with several major problem areas that have been formulated as QPPs. The classical least squares regression problem may be so formulated, as may be linear problems which incorporate uncertainty (variance) into the objective function such as the portfolio problem. Classical economic problems of supply and demand response and diminishing returns to scale may also be cast in a QP framework.

Regression

In least squares regression the mathematical formulation is one of minimizing the squared error term (i.e. least squares). The parameters of the regression equation can be estimated by using a QP formulation which, in certain cases, may be more flexible than a regression approach, since linear constraints can be imposed expressing *a priori* information about the particular problem. With QP, it is possible to estimate regression parameters using either restricted least squares or unrestricted least squares.

The basic unrestricted regression model is

$$\text{Min.} \quad e'Ie \tag{4}$$

$$\text{s.t.} \quad e = Y - X\beta \tag{5}$$

where the β's and e's are the parameters to be estimated, and Y is an $(N \times 1)$ vector of observations on the dependent variable and X is an $(N \times M)$ matrix of observations on the independent variables. The problem is thus quadratic in e subject to linear constraints. The objective function minimizes the summed squared errors by choosing appropriate values of β and e. In order to permit negative and positive β's and e's, a simple well known transformation may be performed; namely, making each parameter the difference between two positive numbers (e.g. $\beta_i = \beta_i^+ - \beta_i^-$, $e_j = e_j^+ - e_j^-$). This unfortunately, makes the tableau rather large and given existing quadratic computational software, a difficult problem to solve.[2] However, the quadratic form permits certain flexibilities such as:

1. *Placing restrictions on* β. If one knew *a priori* that several β's must sum to one (as in constant returns to scale with certain production functions [12]), or that the β_i's were only non-negative (as in the estimation of transition probabilities [21]), then this could be applied directly as a constraint in (4) and (5) above.

2. *Placing restrictions on* e. If one wished to fit a function which only described the upper or lower strata of the data, the e_i's could be restricted to be of only one sign (e.g. $e_i^+ = 0$, indicating only negative deviations or an upper frontier production function [12]). Similarly, to force the regression through a point, the error term could be restricted to be equal to zero.

3. *Weighting the error terms unequally.* In cases where uneven weighting of error terms is desired (such as in Aitkins estimation [7, p. 232], or when primary and secondary information is present), the criterion would be Min $e'Ae$ with A portraying the different weightings.

4. *Assigning special weights to the coefficients.* If one has *a priori* information regarding the values of the β's, the objective function constraints may be modified to reflect penalties for the variables deviating from these levels. (For example, an *a priori* hypothesis might be that $\beta_i = 0{\cdot}8$ and is normally distributed so one could formulate a model to penalize deviations from 0·8).

Basis of QP formulation. Although many of the above flexibilities are available in ordinary least squares analysis, there appear to be several reasons for adopting a QP formulation. First, in modeling the problem in its quadratic form, one has to consider the total problem structure; regression analysis does not require such an explicit formulation, hence, insight into the problem may be lost. Second, when imposing restrictions on the regression, QP provides shadow prices indicating the cost of these restrictions; normally, regression procedures

[2] Actually, if one is interested in unrestricted error terms, the substitution $e = Y - X\beta$ can be made leaving the objective just in terms of β's, saving much space.

and computer codes do not yield this information. Finally, these formulations provide an alternative solution procedure to the user when a proper generalized regression package is not available.

Regression example. Given the data in Table 1, we wish to use QP to estimate the parameters of a linear function relating TV sales to price and production.

TABLE 1. REGRESSION DATA RELATING TV SALES TO PRICE AND PRODUCTION

Observation	Sale of TV's	Price of TV's	Production of TV's
1	100	10	100
2	50	15	100
3	130	13	70
4	200	13	150
5	170	15	200

Assume we have *a priori* knowledge that the price of TV's has less of an effect on sales than does the production of TV's. The QP formulation is

$$\text{Min.} \quad \sum_{i=1}^{5} e_i^{-2} + \sum_{i=1}^{5} e_i^{+2}$$

$$\text{s.t.} \quad -e_1^+ + e_1^- - \beta_0^+ + \beta_0^- - 10\beta_1^+ + 10\beta_1^- - 100\beta_2^+ + 100\beta_2^- = 100$$

$$-e_2^+ + e_2^- - \beta_0^+ - \beta_0^- - 15\beta_1^+ + 15\beta_1^- - 100\beta_2^+ + 100\beta_2^- = 50$$

$$\vdots$$

$$\beta_1^+ - \beta_1^- \qquad - \beta_2^+ + \beta_2^- \geq 0$$

where β_1 and β_2 are the unknown regression parameters associated with price and production respectively. The solution is

$$\begin{aligned}
e_1 &= e_1^+ (=0) - e_1^- (=11{\cdot}24) = -11{\cdot}24 \\
e_2 &= e_2^+ (=0) - e_2^- (=64{\cdot}69) = -64{\cdot}69 \\
e_3 &= e_3^+ (=37{\cdot}38) - e_3^- (=0) = 37{\cdot}38 \\
e_4 &= e_4^+ (=52{\cdot}21) - e_4^- (=0) = 52{\cdot}21 \\
e_5 &= e_5^+ (=0) - e_5^- (=13{\cdot}66) = -13{\cdot}66 \\
\beta_0 &= \beta_0^+ (=35{\cdot}37) - \beta_0^- (=0) = 35{\cdot}37 \\
\beta_1 &= \beta_1^+ (=0{\cdot}6897) - \beta_1^- (=0) = 0{\cdot}6897 \\
\beta_2 &= \beta_2^+ (=0{\cdot}6897) - \beta_2^- (=0) = 0{\cdot}6897.
\end{aligned}$$

The total squared error is 8620·36 and the constraint shadow prices are:

constraint 1	22·48
2	129·38
3	−74·76
4	−104·42
5	27·32
6	245·96.

Vol. 5
1977

The careful reader will note that the shadow prices on the first five constraints are twice the appropriate error terms.[3] An interesting shadow price obtained from the dual formulation of this problem relates to our restriction that β_1 is equal to or greater than β_2. The constraint was binding and the marginal revenue (change in error) for relaxing it is 245·96, clearly demonstrating the cost of this restriction.

Mean-variance (E-V) analysis

Many models have been formulated where the tradeoff between expected return and variance or semi-variance (measures of uncertainty) has been an important consideration. The original work on this topic was done by Markowitz [15] in analyzing stock portfolios. E-V analysis is applicable to many diverse decision problems, whereby a choice must be made among portfolios with uncertain outcomes. These portfolios may consist of such varied items as stocks [18] or agricultural crops [17].

E-V analysis generally consists of two phases: generating the so-called efficient frontier, and selecting an optimal portfolio from this set of efficient portfolios consistent with the decision maker's utility function. In E-V analysis only the first and second moments of the outcome distribution, namely, mean and variance are considered. This presumes that some set of the following three conditions hold: (1) the outcome distribution is completely specified by its first and second moments (as is the case with normally distributed outcomes); or (2) the effect of any moments of the outcome distribution beyond the first two are insignificant; and/or (3) the decision maker's utility function in the decision problem contains terms involving only mean and variance [8, p. 101, 169–178].[4] A set of feasible portfolios and their outcomes can thus be represented in mean-variance (E-V) space. If the decision maker has a risk averse quadratic utility function, i.e.

$$U(X) = X - b\,X^2$$

where U is utility of the outcome X, and b is the risk aversion coefficient, an efficient portfolio is one that has the minimum variance for a given expected return, in the set of all feasible portfolios (Fig. 1). Thus, the set of efficient portfolios defines an upper frontier (A–B), called the E-V frontier. It is from this set of portfolios that a decision maker should choose a portfolio consistent with his utility function.

The efficient frontier is generated from the mean and variance estimates. QP provides an efficient method for deriving the E-V frontier. The formulation can be stated as

[3] Shadow prices are the first derivative of the objection function with respect to the constraint unit. The unit of constraint one is e_1, the derivative of e_1^2 is $2e_1$, hence, the shadow price is as expected.

[4] Any LP can be treated as a classical portfolio problem in the Markowitz sense if the decision variables are considered to be normally distributed random variables.

$$\text{Max.} \quad Z = C'X - \lambda X'\phi X \tag{6}$$

$$\text{s.t.} \quad AX \leq B \tag{7}$$

$$X \geq 0$$

where C is a vector of expected returns, ϕ is the variance–covariance matrix of returns, and λ is a risk aversion coefficient. A is a matrix of technical coefficients for the activities, while B is a vector of resource availabilities.

When λ is zero we have a standard LP. However, when λ is not zero we have a QP. By parameterizing λ from zero to positive infinity (when $\lambda = \infty$, we have a minimum variance problem), we generate the E-V frontier. An example of how an E-V frontier is generated in this way is given by Scott and Baker [17]. An alternate QP formulation for generating the E-V frontier is given in [8, p. 175].

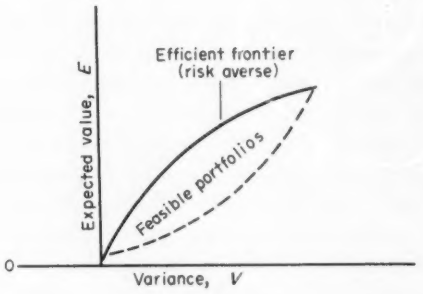


FIG. 1. *E-V frontier.*

Basis of QP formulation. An E-V (i.e. QP) formulation of a constrained linear model, whose decision variables are random variables, has several advantages over a traditional expected value (LP) formulation. The advantages are mainly covered in Sharpe [18]. Principal among the reasons for a QP formulation is to obtain portfolio (or activity) diversification. For example, LP will tend to concentrate an investment in a few stocks due to both the maximization of expected value criterion and the properties of basic feasible solutions.[5] QP, however, will generally result in a diversified portfolio to reduce risk, (e.g. buying multiple stocks), due to the variance–covariance interrelationships of the stocks.

Portfolio example. An investment bank faces the following situation. Three stocks are eligible for investment. Forecasts of performance based on annual rate of return are

[5] In a recent paper Hazell [10] has shown a linear programming formulation which does not have these properties.

Vol. 5
1977

	Stocks		
Forecasts	*I*	*K*	*L*
1	3·00	0·30	4·00
2	1·00	4·50	0·60
3	2·75	1·75	2·75
4	4·50	5·00	1·90
5	3·25	2·75	3·75
Mean forecasted return	2·90	2·86	2·60

The cost of the stocks per share are $I = \$30$, $K = \$45$ and $L = \$27$. Given \$2000 to invest and the restriction that no more than \$1000 can be invested in stocks K and L, the portfolio formulation is as follows:

$$\text{Max.} \quad 2{\cdot}90\,I + 2{\cdot}86\,K + 2{\cdot}60\,L$$

$$-\lambda[I\,K\,L] \begin{bmatrix} 1{\cdot}265 & 0{\cdot}036 & 0{\cdot}640 \\ 0{\cdot}036 & 3{\cdot}013 & -1{\cdot}730 \\ 0{\cdot}640 & -1{\cdot}730 & 1{\cdot}560 \end{bmatrix} \begin{bmatrix} I \\ K \\ L \end{bmatrix}$$

$$\text{s.t.} \quad 30\,I + 45\,K + 27\,L \leq 2000$$
$$45\,K + 27\,L \leq 1000$$
$$I, K, L \geq 0.$$

where the variance–covariance matrix was computed from the above table of forecasted annual returns. Selected solution points are given below and lead to the efficient frontier shown in Fig. 2. S_1, SP_1, S_2, SP_2, are the slack and shadow price from the first and second constraint respectively. At each turning point on the efficient frontier, the shadow price gives the value in return less cost of risk for an increase in funds available.

	Values of λ				
Values of variables	0	0·005	0·028	0·2819	∞
I	66·67	46·40	33·33	0	0
K	0	0	12·80	11·90	0
L	0	22·52	15·69	17·10	0
S_1	0	0	0	1000	2000
SP_1	0·096	0·072	0·70	0	0
S_2	1000	391·80	0	0	1000
SP_2	0	0	0·493	0	0
Variance	5525·55	4754·90	2201·20	139·51	0
Expected value	193·33	193·10	174·10	78·66	0

Supply and demand response

Decision makers often desire information on the equilibrium price and quantity that would exist in a given market. Basic economic theory states that when price changes, so does the quantity demanded. This interaction of price

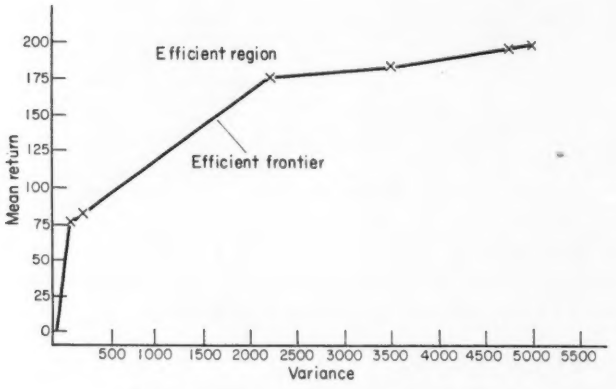


FIG. 2. *E-V frontier for investment problem.*

and quantity can be explicitly formulated as a QP.[6] Samuelson [16] proposed and Dorfman [5] originally formulated the problem as a QP. Judge and Takayama [13] provide an extensive review of the literature. Work in this area has proceeded along two major avenues: (1) classical supply–demand response, and (2) supply–demand response with the inclusion of spatial and temporal considerations [20].

Classical supply and demand systems. Consider the case of an industry which produces N products, using R inputs. The industry is attempting to maximize profits and has the following problem.

$$\text{Max.} \qquad P_y Y - P_x X \tag{9}$$

$$\text{s.t.} \qquad A_1 Y + A_2 X \leq b \tag{10}$$

$$Y, X \geq 0. \tag{11}$$

where P_y is a $(1 \times N)$ vector of product prices;
P_x is a $(1 \times R)$ vector of input prices;
Y is a $(N \times 1)$ vector of product outputs;
X is a $(R \times 1)$ vector of input usages;
A_1, A_2 are $(M \times N)$ and $(M \times R)$ constraint matrices relating inputs to outputs and imposing necessary resource restrictions;
b is a $(M \times 1)$ vector of the resource availabilities.

If supply and demand curves are linear functions expressed as

$$P_y = a_1 - Y'\beta_1$$
$$P_x = a_2 + X'\beta_2$$

[6] This assumption is that the price–quantity response functions (i.e. the demand–supply curves) are linear or piece-wise linear.

where a_1 and a_2 are constant intercept terms, and β_1 and β_2 are matrices of the demand and supply curve slopes, these relationships can be substituted in (9) and (10) to yield the following QP:

$$\text{Max.} \qquad Z = a_1 Y - a_2 X - \alpha Y'\beta_1 Y - \alpha X'\beta_2 X$$

$$\text{s.t.} \qquad A_1 Y + A_2 X \leq b.$$

The parameter α is inserted after the substitution to reflect the degree of monopoly power available to the firm ($\alpha = \frac{1}{2}$ is the case of perfect competition, $\alpha = 1$ is the case of a monopoly) [20].

As with LP, QP equates marginal cost (MC) to marginal revenue (MR) and sets their difference to zero in finding the equilibrium solution. Therefore, when solving quadratic demand and supply problems, with ($\alpha = 1$), the marginal revenue and marginal cost will be equated, and in economic theory this is a monopolistic solution.

If the model is one of perfect competition, then the equilibrium point should occur where the supply and demand curves intersect. By forming an objective function consisting of the integral of the area under the demand and supply curves, Takayama and Judge [20] have shown that if one wants production to reflect perfect competition, α should be equal to $\frac{1}{2}$. This leads to a point where the original supply and demand curves intersect.

Solving the problem of maximizing revenue, ($\alpha = 1$) leads to the intersection of the marginal revenue and marginal cost curves at point p_3q_1 (Fig. 3). For the demand curve, P_y, the price goods are sold for is p_1; for the supply curve, P_x, the cost of the goods is p_2. Clearly, this is a monopolistic situation with the producer making windfall profits. However, if the problem is solved with $\alpha = \frac{1}{2}$, the marginal revenue and marginal cost curves become the actual demand and supply curves and the optimal point is p_3q_2, the purely competitive solution.

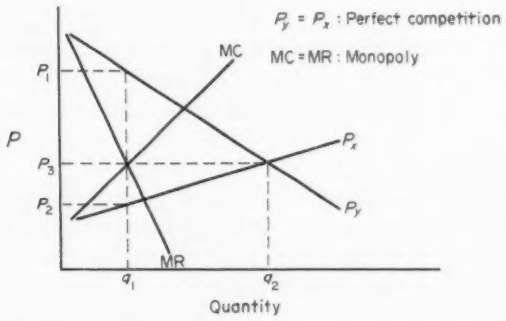


FIG. 3. *Supply–demand response functions: perfect competition vs. monopolistic behavior.*

Basis of QP formulation. The supply–demand formulations allow price and quantity to vary within the solution according to a specified function. The usual LP analysis requires fixed prices. These formulations have been used extensively in the economics literature [13].

Primal information from the above formulations involve the allocation of resources and final products to uses. Dual information is given on the prices for resources and the prices of end products. In the spatial or temporal forms, prices are reported with regional or time differences [20].

Supply–demand example. A firm in Florida producing oranges has two uses for the oranges it produces. One is to convert them into juice, the other is sell fresh oranges. The firm has a supply price (P_S) related to supply quantity (Q_S) as follows:

$$P_S = 3 + 0{\cdot}0005Q_S.$$

Demand price for orange juice (P_J) is related to demand quantity of juice (Q_J) and demand quantity of fresh oranges (Q_F) by

$$P_J = 7{\cdot}5 - 0{\cdot}005Q_J - 0{\cdot}0002Q_F.$$

The price of fresh oranges (P_F) is related to its quantity by

$$P_F = 4 - 0{\cdot}0002Q_J - 0{\cdot}001Q_F.$$

Vol. 5
1977

Assuming a purely competitive market where all prices are those received by the producer and that the producer makes 1 unit of juice from 2 oranges, the following formulation results:

$$\text{Max } P_JQ_J + P_FQ_F - P_SQ_S - 1/2Q'\phi Q$$

$$\text{or Max } 7{\cdot}5Q_J + 4Q_F - 3Q_S - 1/2[Q_JQ_FQ_S] \begin{bmatrix} 0{\cdot}005 & 0{\cdot}0002 & 0 \\ 0{\cdot}0002 & 0{\cdot}001 & 0 \\ 0 & 0 & 0{\cdot}0005 \end{bmatrix} \begin{bmatrix} Q_J \\ Q_F \\ Q_S \end{bmatrix}$$

$$\text{s.t.} \quad 2Q_S + Q_F - Q_S \leq 0.$$

The solution to this problem is: Supply 806 oranges (Q_S) of which 332 are processed into juice yielding 116 units of juice (Q_J), and 574 are sold fresh (Q_F). The per unit price of the products are: fresh oranges (P_F) equal to 3·403, orange juice (P_J) equal to 6·085, supply oranges (P_S) equal to 3·403; the equilibrium price of oranges (= 3·403) is also given by the constraint shadow price.

Diminishing returns to scale

Economists have long treated the problem of diminishing returns to scale. A QP formulation of this problem is straightforward. The objective function involves maximization of returns less the cost of production which is an increasing function of quantity. Mathematically:

$$\text{Max } X'P - X'(C + \phi X) \tag{12}$$

$$\text{s.t.} \quad AX \leq b \tag{13}$$

where p is price;
$C + \phi X$ is the production cost given x units are produced;
$AX \leq b$ are fixed resource constraints.

Using this formulation one derives information similar to the supply–demand response problem above where the cost of production and allocation of production are obtained.

Example. A firm which produces widgets sells them for \$20.00 each, and wishes to produce no more than 200 widgets. Producing the first widget costs \$1.00. After that, costs are an additional 0·00025 dollars. So the cost of a widget when the X_0th widget is made is $1 + 0{\cdot}00025\ X_w$ widgets. The QP formulation is:

$$\text{Max} \quad 20X_w - (1 + 0{\cdot}00025X_w)X_w$$
$$\text{s.t.} \quad X_w \leq 200.$$

The solution is to produce 200 widgets (X_w) where the firm makes \$3790 at the current production cost of \$1.05, with the shadow price on the constraint equal to \$18.95 (which is revenue less current production cost).

POSSIBLE ADDITIONAL FORMULATIONS

As previously indicated, QP formulations adequately portray first order interactions within objective function variables. Such items as falling demand, rising supplies, increasing cost, decreasing revenue, increasing variance, covariance, and first order interaction effects are all possible as demonstrated in the above formulations. These notions can be extended to include other formulations of problems in the management and decision sciences.

First, consider the problem of multiple objectives. Multiple objective programming problems have been extensively treated by Lee [14] and others. However, in these treatments, goals are assumed independent, i.e. satisfaction of one does not affect satisfaction of another. Consider an extreme example where a community is to expand its industrial base or to expand its pollution treatment facilities. One could easily see that if this community satisfied its industrial goal, the demand for pollution treatment would also be greater; clearly a violation of the independence assumption. QP would allow a more realistic formulation and simultaneous analysis of this problem by not assuming independence; that is, by considering the interactions or tradeoffs. Similarly, in the risky case, variances and covariances between goals could be portrayed.

A second promising area of examination arises from approximation theory. Quadratic functions (in the form of second order Taylor series expansions) have long been used to approximate nonlinear functions. Certainly, the application of quadratic approximations to problems with nonlinear objective functions (subject to linear constraints) could lead to satisfactory formulations.[7]

[7] This has long been done in the so-called unconstrained quadratically convergent nonlinear programming algorithms.

Another potential area for examination is the case where quadratics appear in the constraints. Townsley [22] and Chen [3] among others have been able to reformulate such problems with the quadratic in the objective function. For example, problems arising in chance constrained programming [2] have duals which structurally resemble portfolio problems. In fact Wagner [23], in his discussion of chance constrained formulations, states that the probability setting should be considered simultaneously with optimization.

A final area which can be explored is not at all new, but simply combinations of the above formulations. Hazell and Scandizzio [11] recently proposed a problem which involves both uncertainty and falling demand. Certainly composite quadratic formulations may be proposed dealing with several types of interactions simultaneously.

CONCLUSIONS AND LIMITATIONS

QP allows objective function first order interactions. Hence, quadratic formulations are potentially relevant whenever variables in the objective function are interrelated. Most of the problems formulated to date have been 'natural' quadratics. Obviously, there are many other problems which can be so formulated yielding more accurate and realistic formulations. This paper has attempted to summarize some of the natural formulations and suggest some possible future formulations.

Vol. 5
1977

The quadratic formulations in this paper are examples of QP applications in the management and decision sciences. The authors think that examples like those above should supplement the theory in order to achieve balanced classroom presentation and expanded practical application.

Many LP approximations have been proposed for the problems discussed in this paper (e.g. Hazell [10], Hartley and Sielken [9], and Duloy and Norton [6]), so a quadratic formulation is not always necessary. However, to be technically correct, these problems should be formulated conceptually as QPs. Researchers and practitioners alike should consider correct formulation of the problem before worrying about solution difficulties.

The limitations of QP formulations (stated below) are rather straightforward and in fact seem to be easily dealt with.

1. *Size of problem.* Although the capability for large scale QP software exists, generally available QP algorithms are still limited; recently the capabilities for about 400 variables and 200 constraints has been realized [4]. In a recent thesis, Stoecker [19] reported solution of even larger problems. However, if management and decision scientists promote the need for QP solvers, more efficient codes could indeed be rapidly developed.

2. *Data.* Objections to QP formulations mainly relate to data estimation capabilities. However, all applications suffer from such data deficiencies, and

most applications assume away these problems. Perhaps recognition of the capabilities of quadratic programming in conjunction with modern data systems will alleviate these objections.

REFERENCES

1. Boot JCG (1964) *Quadratic Programming*. North Holland, Amsterdam.
2. Charnes A and Cooper WW (1959) Chance-constrained programming. *Mgmt Sci.* **6,** 73–79.
3. Chen J (1973) Feed formulations with a probabilistic constraint. *Am. J. Agric. Econ.* **55,** 175–184.
4. Cutler L and Pass DS (1971) A computer program for quadratic mathematical models to be used for aircraft, design, and other applications involving linear constraints. Report prepared for U.S.A.F. Project Rand, Rand Corp., Project No. R-516-PR.
5. Dorfman R (1951) *Applications of Linear Programming to the Theory of the Firm*. Univ. of California Press, Berkeley.
6. Duloy JH and Norton RD (1975) Prices and incomes in linear programming models. *Am. J. Agric. Econ.* **57,** 591–600.
7. Goldberger AS (1964) *Economic Theory*. John Wiley, New York.
8. Halter AN and Dean RW (1971) *Decisions Under Uncertainty with Research Applications*. South-Western, Cincinnati, Ohio.
9. Hartley HO and Sielken LL (1973) Two linear programming algorithms for unbiased estimation of linear models. *J. Am. Statist. Ass.* **68,** 639–641.
10. Hazell PBR (1971) A linear alternative to quadratic and semivariance programming for farm planning under uncertainty. *Am. J. Agric. Econ.* **53,** 53–62.
11. Hazell PBR and Scandizzo PL (1974) Competitive demand structures under risk in agricultural linear programming models. *Am. J. Agric. Econ.* **56,** 235–243.
12. Johansen L (1972) *Production Functions*. North-Holland, Amsterdam.
13. Judge GG and Takayama T (1973) *Studies in Economic Planning Over Space and Time*. North Holland, Amsterdam.
14. Lee S and Moore L (1973) Optimizing transportation problems with multiple objectives. *AIIE Trans* 5.
15. Markowitz HM (1959) *Portfolio Selection, Efficient Diversification of Investments*. Cowles Foundation Monograph 16. John Wiley, New York.
16. Samuelson PA (1952) Spatial price equilibrium and linear programming. *Am. Econ. Rev.* **42,** 283–304.
17. Scott JT and Baker CB (1972) A practical way to select an optimum farm plan under risk. *Am. J. Agric. Econ.* **54,** 657–661.
18. Sharpe WF (1970) *Portfolio Theory and Capital Markets*. McGraw-Hill, New York.
19. Stoecker AL (1974) A quadratic programming model of United States agriculture in 1980: theory and application. Ph.D. thesis, Iowa State University.
20. Takayama T and Judge GG (1971) *Spatial and Temporal Price and Allocation Models*. North-Holland, London.
21. Theil H and Ray G (1966) A quadratic programming approach to the estimation of transition probabilities. *Mgmt Sci.* **12,** 714–722.
22. Townsley R (1968) Derivation of optimal livestock rations using quadratic programming. *J. Agric. Econ.* **19,** 347–354.
23. Wagner HM (1975) *Principles of Operations Research*, 2nd ed., Prentice-Hall, Englewood Cliffs, NJ.

ADDRESS FOR CORRESPONDENCE: *Professor H Moskowitz, Krannert Graduate School of Management, Krannert Building, West Lafayette, Indiana* 47907, *USA.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

A Multicriteria Approach to Strategic Planning: An Application in Inventory Control

ROBERT C CARLSON
Stanford University, California

HANS H THERP
McKinsey and Company, Copenhagen, Denmark

(Received May 1976; in revised form August 1976)

A new planning method is proposed whereby a manager can observe the strategic consequences of his tactical decisions before those decisions are implemented. These consequences are presented as goal outcomes or values of criteria previously specified as being important by the decision maker. The methodology involves the embedding of relatively simple tactical optimization models within a larger and more complex descriptive model. The results obtained in applying this method to an electronics manufacturer are presented. Of particular interest is the comparison among three tactical inventory models.

1. INTRODUCTION

INVENTORY control is a total problem where management is interested in utilizing its resources in the best way—this is the strategic problem. But an inventory control system has many subsystems, items to be controlled in the best way—these are the tactical problems. Obviously, strategic and tactical problems are interdependent, and one should not be considered without concern for the other. This fact is often neglected in systems today because no relationship has been established between the tactical and the strategic decisions. Tactical decisions are made without considering the strategic consequences; but once the tactical decisions have been made, the strategic performance is implicitly determined.

Numerous inventory control models exist which minimize the costs incurred for each item, considering only that item. Based on estimates of the cost coefficients involved, this gives a solution to the tactical problem of how to control each item. When such models are used in practice, the resulting strategic solutions are often far from satisfactory [6]. The real problem is the selection of proper costs. The cost coefficients are estimates of marginal costs which are

difficult to obtain even from historical data. Estimates of future marginal costs will normally require some knowledge about the resulting outcome levels; but these are strategic consequences which are not considered in traditional inventory control methods.

2. ANALYZING STRATEGIC CONSEQUENCES

The total costs of an inventory system arise because capital is tied up in inventory and facilities, because manpower is needed to obtain and maintain the inventory, and because a shortage situation may cause paper-work, lost sales, or delayed production. Commonly the performance of an inventory system is expressed by three outcomes—average inventory level, number of orders processed per year, and the level of backorders. These three outcomes can be considered goals of the inventory system, and their aggregate values are the strategic consequences of the inventory policy used.

The tactical decision variables used in many inventory systems are an order point and an order quantity for each item. When the inventory level of an item is reviewed, a replenishment order is processed when the amount in stock and on order is less than the order point for that item. Many inventory control models suggest order points and order quantities which are optimal with respect to the criterion function of the model. Generally, however, the criterion function is a simplification of the true objectives of the inventory system. It includes estimates of cost coefficients; and the outcomes are expressed as a simplified function of the tactical decisions and the item parameters to facilitate finding 'optimal' decisions.

This paper suggests the use of 'near optimal' tactical decisions derived from a traditional inventory model, but only after analyzing the implied strategic consequence. By changing the ratios between the cost co-efficients used in the inventory model, different tactical decisions will be 'optimal'. It is assumed that the cost ratios are common for all items.

The strategic consequences are the aggregate values of the outcomes, and they are calculated for each set of tactical decisions. The calculation of the strategic consequences is the prediction of the strategic performance of the inventory system if a particular set of tactical decisions is used. The contribution from each item to the strategic consequences must be calculated more accurately than can be done using the approximate relationships used in the tactical inventory models. This is essential since a practical inventory system almost always is more complex than assumed in the inventory model, and the value of analyzing the strategic consequences depends on the accuracy with which they can be calculated. Often a simulation model is used [2, 3].

The inventory model enables us to calculate the strategic consequences corresponding to a set of cost ratios. Each set of strategic consequences is a

point in outcome space, and the points corresponding to different cost ratios constitute the 'efficient surface' [1,2] for that particular model. The efficient surface informs the decision maker about the alternatives which the considered inventory model allows. The strategic decision is made by selecting a point on the efficient surface [5].

Through the efficient surface, managers obtain insight into the trade-offs that exist between the inventory level, the number of orders processed, and the level of backorders. Through investigation and discussion of alternative combinations of these outcomes, the managers can weight the outcomes and agree upon a preferred point on the efficient surface. In this way managers are given the possibility of direct involvement in making strategic decisions and hence direct influence on strategic consequences. The information, experience, and intuition of the managers can be utilized in the strategic decision instead of only in the evaluation of the past performance of the inventory system.

After the selection of the preferred point on the efficient surface, the problem is to identify those tactical decisions which will lead to the particular combination of strategic outcomes defined by the point. Since each point on the efficient surface corresponds to a set of cost ratios, we can find the set of cost ratios which corresponds to the preferred point selected above. This set of cost ratios should then be used in the inventory model, and the optimal tactical decisions identified will correspond to the preferred point on the efficient surface. When these tactical decisions are used in the coming periods, the strategic decisions are implemented.

3. A REAL-WORLD INVENTORY CONTROL APPLICATION

Our method for strategic decision making was applied to a real-world inventory system consisting of approx 7000 parts that are used in the production of measuring instruments. Production of the instruments was in batches, called runs, according to a production schedule. The sales rate for each instrument fluctuated, and for all practical purposes we could assume these sales rates to be independent. Because the production schedule was changed every month to reflect realized sales and forecast changes, the usage of parts deviated from the planned level. These deviations were of course not strictly independent; but the practical situation was such that many changes in the production schedule occurred every month, and furthermore most parts went into several instruments. We therefore made the assumption that the usage rates of the parts were normally distributed and independent.

The parts were purchased on an order point/order quantity basis, and we could use the framework of the strategic planning method to evaluate alternative

purchasing rules for the inventory system. The inventory performance was measured in terms of the following three goals:
average total inventory;
number of replenishment orders per year;
number of backorders at any point in time.

An inventory analysis was undertaken to find the impact of the purchasing rules used on the performance of the inventory system. The following two studies were performed: (1) comparison between the theoretical performance of the current purchasing rules and purchasing rules derived from two different inventory models; (2) analysis of the impact of forecast errors on the performance of the inventory system.

Since the calculation of the tactical decisions for a part involves an optimization algorithm, the computations can be considerable. We therefore performed the consequence calculations for a sample of the parts and extrapolated the results to obtain estimates of the strategic consequences for the total inventory.

The inventory considered here consists of 7300 parts classified as A, B, C and D parts, shown in Table 1.

TABLE 1. A, B, C, D CLASSIFICATION OF PARTS

Class	% of $ usage	% of inventory $	Number of parts	Sample size
A	77	61	720	100
B	17	23	1130	100
C	5	12	1530	100
D	1	4	3920	100

First we calculated the three consequences for the 100 A parts. Each of these three consequences was then multiplied by the ratio between the total forecast for all A parts and the forecast for the 100 parts in the sample. In this way we estimated strategic consequences for the inventory of all A parts. We repeated these calculations for the other three classes and added the consequences for all four classes.

In addition to reducing the computations involved, sampling greatly reduces the amount of part data required by the method. For the 400 parts in the samples, the current list prices and leadtimes were retrieved from a data file. The number of monthly pulls and the size of each pull were obtained from a file containing past transactions for the 400 parts covering a 6-month period. The forecast of the usage for each part was based on the part explosion of scheduled production for the 6-month period. This forecast was revised every month, but for simplicity, a single 6-month forecast was used. From the data on monthly pulls, we estimated the average pull size, the mean actual usage, and the standard deviation of the usage rate for each part. The existing purchasing rules specified

Vol. 5
1977

a set of tactical decisions for all parts. Different order points and order quantities were used depending on the classification of a part, the usage rate, and the lead-time (Table 2). These tactical decisions were not based on any inventory model, but on intuitive rules adjusted over the years using feedback from actual inventory performance.

The theoretical performance of these rules was then calculated for these order points and order quantities for the samples of A, B, C and D parts. The performance of the total inventory was estimated using the sampling procedure previously described. The reorder quantities (EOQ's) were computed through the use of the conventional square root EOQ formula.

TABLE 2. CURRENT PURCHASING RULES

The order point (OP) for each part is determined by OP = (safety stock + leadtime) × (weekly usage) where safety stock (in weeks of coverage) is determined by

Leadtime	Safety stock
0–8 weeks	2 weeks
8–14 weeks	3 weeks
14–20 weeks	4 weeks
above 20 weeks	5 weeks

The order quantity is determined by

Class	Order quantity
A	6 weeks usage
B	8 weeks usage
C	20 weeks usage
D	32 weeks usage

The safety stock was taken as [7], $P \times \sigma_l$, where P is a constant determined by a policy decision and σ_l the standard deviation of demand during the leadtime. Using this safety stock calculation for all parts theoretically implies the same chance of stockouts for all parts during each order cycle. In our case we assumed that σ_l was the square root of the mean demand during the leadtime [4].

Under the assumptions of normally distributed, independent usage rates, our inventory model optimizes the combination of the three consequences used in evaluating the inventory performance. We then used some approximations [3, 4] for the case of continuous review and independent, normally distributed usage rate.

4. RESULTS OF ANALYSIS

Figure 1 shows the possible combinations of the strategic consequences for $P = 2$ and $P = 3{\cdot}7$ respectively. Figure 2 shows the results for the corresponding

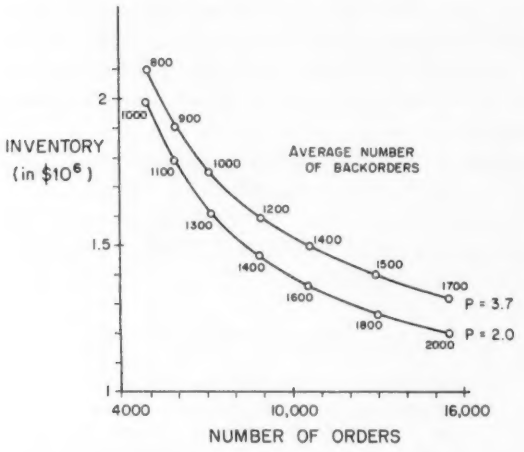


FIG. 1. *Efficient surface for EOQ model (using actual forecast).*

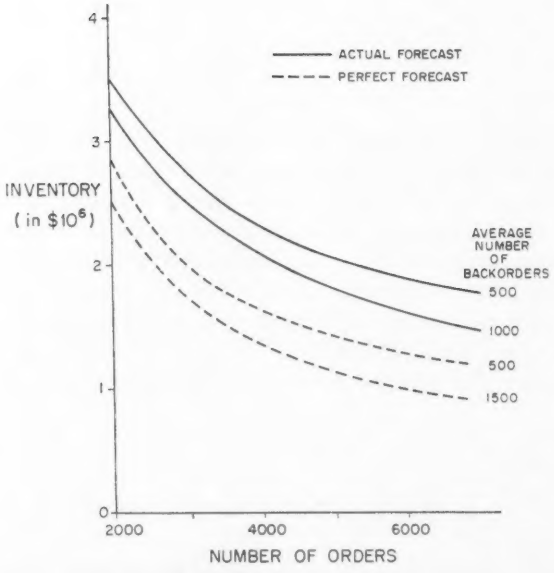


FIG. 2. *Efficient surface for stochastic model.*

stochastic inventory model. It is obvious that the three inventory control techniques perform very differently. As we would expect, the stochastic inventory model can always find solutions that dominate those of the EOQ model, and the EOQ model in turn can produce solutions that dominate the consequences of the current buying rules. A closer comparison is made in Table 3.

We then investigated the impact of a general forecast error on the goals of the inventory system. If the aggregated forecast is too high, the actual inventory level will be higher than expected; and if the forecast is too low, the number of backorders will be higher than expected. We calculated the inventory performance under two conditions: the first as in the foregoing analysis (order points and order quantities based on the forecast, while the inventory performance on the actual average usage rate of each part), the second relies on a perfect forecast (i.e. the forecast and actual demand coincide). The difference in the three strategic consequences between these two calculations can be attributed to the forecast errors.

TABLE 3. COMPARISON OF THE THREE INVENTORY TECHNIQUES

For number of orders per year = 15,000 and expected number of backorders = 1600

Technique	Inventory
Current	$1,730,000
EOQ ($P = 4$)	$1,370,000
Stochastic	$1,150,000

For inventory level = $1,730,000 and expected number of backorders = 1100

Technique	Number of orders per year
EOQ	7000
Stochastic	5100

For inventory level = $1,730,000 and number of orders per year = 7000

Technique	Expected number of backorders
EOQ	1100
Stochastic	800

In the example used, the forecast on the aggregate was 33% too high. The results for the two models are shown in Table 4.

TABLE 4. EFFECT OF FORECAST ERRORS

	Forecast	Inventory	Number of orders	Expected number of backorders
Model 1: Current purchasing rules	Actual	$1,730,500	15,500	1630
	Perfect	$931,700	17,800	1425
Model 2: Stochastic inventory model	Actual	$1,730,000	5100	1100
	Perfect	$1,080,000	5100	1100

A warning

A number of factors exist in practice which have not been included in this analysis. They all have an impact on the performance and the calculations of the performance. It is important to note that these factors not only affect the absolute accuracy of the consequence calculation, but have less influence on the relative performance of the three techniques compared. Some of the factors are listed here.

1. Overriding of purchasing rules based on detailed knowledge about the production schedule often occurs.
2. The use of split shipments was not considered.
3. No consideration was given to quantity discounts in the models.
4. Minimum shipments are required by some vendors.
5. The revision of forecast every month was not included.
6. Six month usage data may not be enough to estimate the standard deviation reliably.
7. The usage rate may not be normally distributed and independent.

5. CONCLUSIONS

The proposed method of calculating the efficient surface for an inventory system is a tool which enables managers to deal with the strategic decisions for inventory systems. The efficient surface tells the managers about alternative combinations of the three outcomes—average inventory level, number of orders per year, and the level of backorders. They can relate these outcomes to the overall objectives and operating restrictions of the company. The method will not calculate one magic optimal solution, but it will leave the final strategic decision to the managers and thus make it possible to include their experience and intuition. The method will prove particularly useful in adjusting to changes in the environment which affect the inventory system such as the supply of

capital, leadtimes, and cost changes. The method can be used with any existing or future tactical inventory models, and strategic consequences other than those considered here can be included.

REFERENCES

1. BAUM S and CARLSON RC (1974) Multi-goal optimization in managerial science. *Omega* 2(5), 607–623.
2. FEENEY GJ (1955) A basis for strategic decisions on inventory control operations. *Mgmt Sci.* 2(1), 69–82.
3. GUNGE T and THERP HH (1972) *Ressourceallokering i Lagerstyring*. Thesis from the Institute of Mathematical Statistics and Operations Research at the Technical University of Denmark.
4. HADLEY G and WHITIN TM (1963) *Analysis of Inventory Systems*. Prentice-Hall, Englewood Cliffs.
5. HANSSMANN F (1962) *Operations Research in Production and Inventory Control*. Wiley, New York.
6. HARTY JD, PLOSSL GW and WIGHT OW (1963) *Management of lot-size inventories*. American Production and Inventory Control Society, Special Report.
7. WIGHT OW (1974) *Production and Inventory Management in the Computer Age*. Cahners Publishing Company.

ADDRESS FOR CORRESPONDENCE: *Professor RC Carlson, Department of Industrial Engineering, Stanford University, School of Engineering, Stanford, California* 94305, *USA.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

Spatial Economies in Distribution Transports

R LAULAJAINEN
Helsinki School of Economics

(*Received June 1976; in revised form August 1976*)

Unit costs of distribution transport (depot → customer) are studied using a simulator. Petroleum companies with four spatial distribution parameters—size (= market share), market area, depot pattern and customer structure—are used as observations. Relevant combinations of the parameters are evaluated with real world data. The unexpectedly small differences between alternatives appear to be due to the dominance of a single urban agglomeration.

l. 5
77

PROBLEM

IT IS a well established economic doctrine that company unit costs decline as a function of increased scale of operations. The decrease is customarily attributed to production technology and most empirical evidence is related to it. Other aspects such as financing and marketing, with the subdisciplines of advertising, sales promotion, R&D and physical distribution, have received far less attention in this context. Nevertheless, the possibilities offered for savings, e.g. by changes in the spatial intensity of distribution, are obvious. Intensity need not, of course, be the only spatial dimension of unit costs of distribution.

In this paper an attempt is made to measure the significance of various spatial distribution strategies for the unit costs of distribution transports.[1] The attempt is based on data on shipments, orders and delivery costs from actual operations.

The most crucial spatial aspects of distribution transports, apart from the intensity of operations, are probably the location and extent of delivery areas (= market area), the number and location of depots and the location and size of customers within delivery areas. Customer size as a spatial concept depends on the observation that, on average, the largest orders with lowest unit costs originate from the largest customers in locations with the largest sales potential.

[1] Distribution transports are defined as shipments from depot to customers. In the narrow meaning they cover only transports in which the vehicle load is, in principle, delivered to several destinations.

The delivery areas and the depots depend on specific decisions made by the company. Specific decisions may also be made as to the size and 'quality' of customers even if there is an element of chance in the actual outcome of sales activity. Consequently, a suitable combination of delivery areas, depots and *a priori* customers should result in a minimization of company unit costs. Because of the hypothetical internal economies of scale, the combination is viewed against the company size.

This approach tells little about company profits, of course. It is possible, for example, that the strategies adopted by several companies may be directed towards the same locations resulting in price competition and decline of profits. The partial nature of the discussion is further stressed by the fact that the study is restricted to a very limited area. This restriction is imposed purely for practical reasons of data collection and analysis. It has the unfortunate result that many of the delivery areas are truncated.

Because of the partial nature of the discussion, there appears to be little reason to seek out a minimum cost strategy. Instead, the unit costs of actual distribution strategies or alternative strategies worthy of serious consideration are discussed. The latter in addition to the random element in customer structure, implies that actual accounting data from existing companies has to be re-arranged. This is done using a *Monte Carlo* simulator, the primary analytical device of the study.

Vol. 5
1977

SIMULATOR

The *Monte Carlo* simulator is used here to produce unit cost data on companies using various distribution strategies. Due to heavy resource requirements, the simulator has not been tested with historical data but relies solely on the critical assessments of the staff of distribution companies. The simulator is static in the sense that an order set is generated *in toto* before it is delivered and that it is delivered completely before a new set is generated (Fig. 1). Spatial location in the Study Area is approximated by a square net. The simulator's road net is based on this square net.

The flowchart of the simulator needs comment: in *the customer generator* the probability of the company having customers in a square equals the total sales volume of the square for all companies. The company market share must fall within certain bounds tied to the total market share of the company. The bounds get narrower as the total sales volume for all companies within the square grows. In squares with very small sales the indivisibility of customers is accounted for by assuming a market share of 100% in every n^{th} square selected. Two types of customers are used. For one (service stations) a certain threshold size in a square is required.

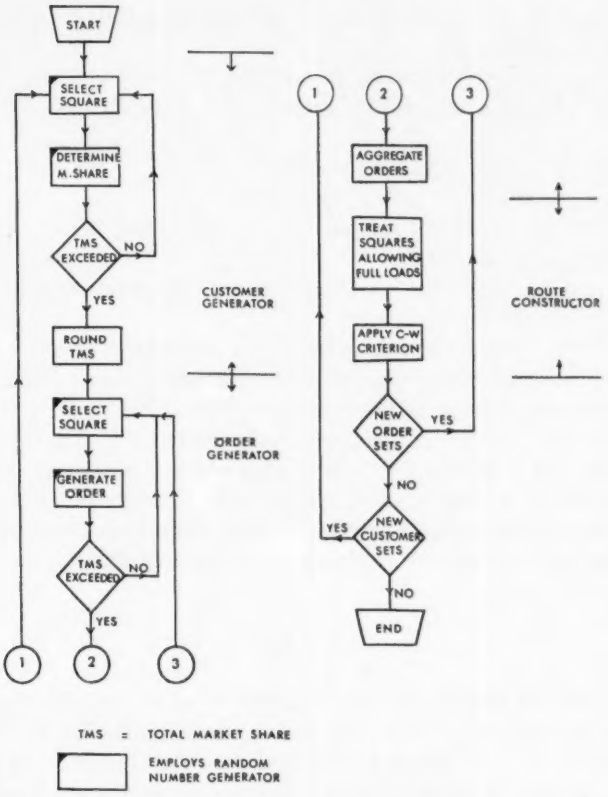


FIG. 1. *Flowchart of the simulator.*

In *the order generator* the total market share allocated by the customer generator is rounded so that the subsequent order volume is an exact multiple of the loading capacity of a large truck (see below). The square selection is proportional to the company sales potential in the square allocated by the customer generator. The order size is a function of the total sales volume of the square for all companies and the type of customer served. All orders originating from a particular square are aggregated to a whole. It is assumed that these orders are of equal size. Variation in order size results thereafter only from the variation between squares (with varying shares of the type of customers). All orders are assumed to be multiples of 1·5 m^3. This simplification obviates the need to consider the vehicle loading problem with different product qualities.

For every company five customer sets and for every customer set five order sets are generated. All input values for a particular customer set and order set are identical except for the seed numbers of random number generators. Every

overall market share. It is sufficient to have a large share in a limited area with reasonable sales potential, as exemplified by some of the small companies in the late 1960's. So, the relevant question is more the absolute magnitude of the internal economies than their qualitative existence. The answer serves as a clue to the intriguing problem of freedom of entry to the market.

The selection of the depots used also appears obvious in the qualitative sense. A limited scale of operations was not sufficient to cover the fixed costs of running a depot in Jyväskylä or in Kokkola. Again, the relevant question is the quantitative impact of deleting these locations and, in the case of Jyväskylä, of variable depot throughput. The first question becomes relevant e.g. for a growing company which up to now has operated only in the southern part of Central Finland from a depot in Helsinki. For a company having a depot in Jyväskylä the Porvoo refinery offered a low-cost alternative source for between depot transfers.[3]

The outcome of customer structure is straightforward in the sense that, on average, service stations place larger orders than 'other customers', thus implying a lower unit cost. Unfortunately they usually require some kind of financial involvement by the supply company, which probably offsets the saving in distribution costs. The choice whether to supply the service stations or not applied only to a small company, however. They could not be neglected by larger companies.

Vol. 5
1977

This discussion is consolidated into five types of companies to be simulated (Table 1). They represent the range of spatial distribution strategies applied or

TABLE 1. ALTERNATIVE DISTRIBUTION STRATEGIES

Company type	Market input %	share output[1] %	Thruput of Jlä	Market area	Depots	Customers[2]	Unit cost[3] (Fmk/m³)
1	10	8·4	0	Central Finland	Hki, Kla	1, 2	25·97
	15	13·9	0	Central Finland	Hki, Kla	1, 2	25·72
	20	19·5	0	Central Finland	Hki, Kla	1, 2	25·49
	25	23·9	0	Central Finland	Hki, Kla	1, 2	25·05
	30	27·8	0	Central Finland	Hki, Kla	1, 2	25·04
2	5	2·8	0	Jlä with surr.	Hki	2	25·22
3	10	8·4	0	Central Finland	Hki, Kla	2	26·39
4	10	8·4	0	Central Finland	Hki	2	29·14
5	20	19·5	100	Central Finland	Hki, Kla, Jlä	1, 2	28·67
	20	16·7	140	Central Finland	Hki, Kla, Jlä	1, 2	28·90
	30	27·8	100	Central Finland	Hki, Kla, Jlä	1, 2	26·41
	30	29·0	140	Central Finland	Hki, Kla, Jlä	1, 2	27·91

[1] Output shares with identical input shares may vary due to rounding by the order generator.
[2] 1 = service stations, 2 = other types of customer.
[3] Only company averages are given.

[3] Because the prices of domestic products are maintained level with those of imports, domestic and imported products can be considered substitutes.

applicable in the early 1970's and shortly before. Type 1 measures the impact of mere company size. Type 2 challenges the undifferentiated economies of scale with a concentrated market area and consequent depot selection. Type 3 compared with type 4 approximates the significance of the Kokkola depot while type 5 compared with the equal-sized companies of type 1 evaluates the Jyväskylä depot. Lastly, type 3 compared with the equal-sized company of type 1, measures the outcome of not supplying service stations. No simulated company is a strict replica of an actual company, of course. However, only those of type 1 with an input market share of 20–30% are completely fictitious, though plausible companies.

Quantitative evaluation of the strategies is based on cost data from delivery transportation and such transfer and storage costs as are necessary to compare the alternatives. By taking the refinery/coastal depot gate as a source, two alternative transport chains are obtained:

1. Porvoo refinery $\xrightarrow{\text{rail}}$ Jyväskylä depot $\xrightarrow{\text{road}}$ customer.
2. Helsinki/Kokkola depot $\xrightarrow{\text{road}}$ customer.

The transportation rates apply for road and rail shipments in 1973. Trucking rates are assumed valid for both owned and rented trucks. Preliminary simulations suggested that to minimize trucking costs only one truck size (large) should be used. Additional sea freight to Kokkola and the capital cost for winter storage there can be safely estimated. At the Jyväskylä depot only costs variable at 2–3 months' notice, for example wages, are taken into account.

The empirical data basis of the customer generator, 1800 deliveries in 1963, is calibrated to correspond approximately to one working day in the early 1970's. Data for the order generator, 2500 orders in 1970 and 1971, are considered valid as such.

The route constructor necessitates advance determination of depot delivery areas. The border between Helsinki and Kokkola is, in principle, equidistant from both depots. To allow for additional sea freight and winter storage, distances from Kokkola are, however, increased by 20 km. The optional Jyväskylä delivery area comprises as many squares as necessary as close to Jyväskylä as possible so that the desired depot throughput is reached.

RESULTS

The spatially undifferentiated scale economies exemplified by type 1 change linearly and are of the order of 4% in the 10–30% range of input market share (Table 1 and Fig. 3). By concentrating operations to the parts of central Finland with the largest sales potential and most 'coastal' location, an overall input market share of 5% by type 2 gives the same unit cost as the undifferentiated company with a share of 20–25%. The overall market share of 5% corresponds

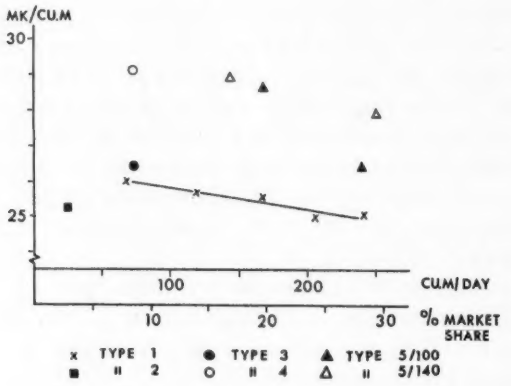


FIG. 3. *Unit cost as a function of the four spatial competition parameters. Only company averages are indicated.*

to a local share of 10–12%. The cost of distribution transport does not seriously deter entry to the petroleum market.

The existence of a depot in Kokkola is a strong asset for a company of type 3 operating all over Central Finland, as the unit cost of type 4 is 10% higher. The higher cost cannot be completely avoided, however. When a small company of type 2 grows larger it must, at some phase, extend its market area. Unless the extension takes place simultaneously on the west coast and in Central Finland there is inevitably a period during which shipments to northern Central Finland are not yet sufficient to support depot operations on the west coast.

The unexpected outcome of the study is the high unit cost of type 5 companies, particularly bearing in mind that no fixed costs for the Jyväskylä depot are included and that alternative strategies of type 1 with 5–10% lower unit costs were not employed in actual operations. The relatively high rail freight to the Jyväskylä depot, in fact two-thirds of the corresponding trucking rate, is the major reason for the cost level in type 5. The hypothetical remedy, scale economy from larger throughput at Jyväskylä, is lost through the costs of delivering to a wider market.

One argument for the existence of type 5 can be found in the length of the planning period. For example, since 1973 the road freight rates have increased compared to rail rates and the cost drawback of the Jyväskylä depot has almost disappeared. The Jyväskylä depot also derives direct benefit from two other factors which are, however, more difficult to express in money terms, i.e. shorter delivery time implying better service to customers and the fact that deliveries are made within the normal working day of truck drivers.

The last question, that of customer structure, proves relatively insignificant

in the sense of distribution unit cost. The unit cost of type 3 is less than 2% higher than that of an equalsized company of type 1.

In general, differences between the alternatives are probably smaller than expected. The major reason seems to be the relative importance of Jyväskylä as a potential market with the result that it is included in all distribution strategies and it clearly dominates them. The dominant role of urban agglomerations, also being a factor elsewhere in Finland, as it is in many other countries as well, the concept of *relative* insignificance of the unit cost of distribution transports to alternative spatial distribution strategies can probably be extended beyond the borders of Central Finland.

CONCLUSION

In this paper an attempt was made to measure the significance of various spatial distribution strategies to the unit cost of distribution transports. The main analytical device was a simulator. Transportation of clean petroleum products in Central Finland in the early 1970's was used as the framework of the simulations and as the empirical data basis.

Differences in unit cost between relevant distribution strategies were at most 10%. Even then the high-cost companies could be shown to represent a transitory phase or to benefit from items not included to the calculations. In particular, the small size of a company was not necessarily a barrier to low-cost distribution transport. The heavy concentration of market potential to a single location was probably the decisive reason for the conclusions and a factor of general relevance.

ACKNOWLEDGEMENT

Acknowledgement is made to the Petroleum Companies operating in Central Finland for supplying the necessary data. The computer staff of The Helsinki School of Economics was of invaluable help in several phases of designing the simulator. Professors Pertti Järvinen, Tampere; Olli Lokki, Helsinki; and Yrjö Seppälä, Helsinki have commented on the general setting of the study. The figures were drawn by Mrs. Eeva Kallioniemi. This particular paper, being part of a larger project, was supported by The Neste Foundation and The Niilo Helander Foundation.

REFERENCES

1. BAIN JS (1965) *Barriers to New Competition.* Harvard University Press, Cambridge.
2. CLARKE G and WRIGHT JW (1964) Scheduling of vehicles from a central depot to a number of delivery points. *Ops Res.* **12**(4), 568–581.

3. DOLL CL (1973) A preliminary simulation study of the delivery problem in an urban area. Centre of Transportation Studies, Working paper No. 10, The University of British Columbia, Vancouver.
4. DUGUID AM and MOSS CG (1971) Reorganizing of sales and distribution system. *Opl Res.* **22,** 77–92.
5. EILON S, WATSON-GANDY CDT and CHRISTOFIDES N (1971) *Distribution Management: Mathematical Modelling and Practical Analysis.* Griffin, London.
6. GRILICHES Z and RINGSTAD V (1971) *Economies of Scale and the Production Function, an Econometric Study of Norwegian Manufacturing Establishment Data.* North-Holland, Amsterdam.
7. LAULAJAINEN R (1973) An entrepreneurial view of transportation. *Geografiska Annaler* **55B**(2), 106–120.
8. SILBERSTON A (1973) Economies of scale in theory and practice. In *Readings in Applied Microeconomics.* (Eds. WAGNER L and BALTAZZIS N). Open University Press, London.
9. SUSSAMS JE (1971) *Efficient Road Transport Scheduling.* Gower, London.

ADDRESS FOR CORRESPONDENCE: *Professor R Laulajainen, Topeliuksenkatu 7A4,* 00250 *Helsinki* 25, *Finland.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

Analysis of a Partially-Observed Queuing System

HOWARD A PLOTKIN
University of Houston, Texas

(*Received March 1976; in revised form August 1976*)

This paper describes the behavior of communication on a land mobile radio channel and suggests a mathematical model to be used in analyzing such communication. Mobile radio is gaining wide recognition as an important business tool, but its growing use brings the problem of congestion. Although the communication on one radio channel by twelve different companies has the structure of a complex queuing system not all of which can be observed, the paper shows that such communication has some of the behavior of a simple queue. This facilitates the analysis of the radio channel and provides an illustration that models of simple queues may offer a helpful means of analyzing those which are more complicated. Once the situation can be analyzed it becomes possible to try to manage it as effectively as possible. In the case of mobile radio the analysis provides a means of measuring the cost of congestion under alternate circumstances.

INTRODUCTION

MANY businesses use the radio to communicate between a main office and personnel operating throughout a city or some other local region. In addition, the units in the field may use the radio to communicate among themselves. In doing so, the radio operates on a channel—a small region of radio spectrum. Since there are more businesses using mobile radio than there are available communication channels, most businesses must share a radio channel with other companies. When sharing, it is necessary that one listens for a few seconds before sending a message, to avoid talking in the middle of someone else's conversation. If the radio is already in use, the message must wait. Delays in business generally have economic consequences. Queuing analysis has been used to evaluate the cost of delays in mining and trucking operations [2, 6], for example. The general issue of costs associated with queuing delays has been discussed elsewhere [4, 5].

The sharing of radio channels which has resulted from the rapid growth in land mobile radio communication has led to a condition of extreme congestion on an increasing number of channels. Such congestion causes message delays which are often undesirably long. With projections of increased utilization of

this type of communication, congestion will intensify. When a company desires to start using mobile radio, a government agency (Federal Communications Commission in the US) assigns it a particular radio channel. In doing so, congestion on that channel, already assigned to others will increase. At what point should they no longer assign additional companies to a particular channel? How can they determine the best channel to assign a new user? How can they determine the effect of assigning the additional company to the radio channel? These questions need to be answered if the radio channels are to be managed in an optimum manner. To answer these questions we must be able to describe the behavior of the communication system.

The process by which the need to communicate arises at any of the main offices or mobile units and by which the messages possibly wait and are eventually transmitted can be viewed as the operation of a queuing system. The system is somewhat complicated, with message generation rates and message lengths different at the various companies. We cannot record the moment in time when someone decides he would like to send a message. In addition, we are unable to observe the number of messages that are waiting, as notes on paper or as ideas in the minds of the people involved. Complicated queuing systems are commonly analyzed by means of computer simulation. In this paper we investigate the ability of the model of a different system to provide analytic means of describing the behavior of this complex system.

Vol. 5
1977

THE INVESTIGATION

Twelve businesses share a UHF radio channel and a repeater on the top of a tall building in downtown Houston. Records of radio salesmen and the Federal Communications Commission indicated sixteen companies assigned to that channel. However, four of the companies were no longer operating there. The radio users in the Business Radio Service, were the following types:

1. electrician;
2. contractor;
3. refrigerator and air conditioner installation;
4. air conditioning and heating service;
5. lawn sprinkler sales and service;
6. delivery service;
7. garbage collection;
8. plumber;
9. heavy equipment sales;
10. construction;
11. tire service;
12. television service.

All were small companies, none having more than six mobile radios.

Each company was visited to obtain an understanding of its use of mobile radio and comments on the experience with radio. No company knew the total number of companies sharing the channel. Estimates varied from 4 to 15. About half the companies knew the names of some of the companies with whom they shared the channel, but the other half did not know the names of any of the others.

The twelve companies had been using mobile radio and sharing their channel for periods ranging from 6 months to 3 years. All but one indicated that the channel is busier now than when they began using it. All but two companies expect their own use of mobile radio to increase in the next few years. It would appear that any problems associated with congestion will get worse.

During a 1 month period the radio channel was monitored and tape-recorded on several days and at different times during the normal work day (8 a.m.–5 p.m.). This permitted as detailed an analysis of all communication and non-communication as was possible. Some judgement was involved in the analysis. For example, it was not always clear whether a period of communication by one company consisted of two or more consecutive messages or one involved message among several personnel. Once the judgement was made, the lengths of the messages and the lengths of periods without communication (idle periods) were timed. The sequence of all messages and idle periods was observed, from which the busy periods could be determined. From these data it was also possible to find the times between successive message completions.

STRUCTURE OF THE ANALYSIS

The radio operation can be viewed as a single server queuing system with messages as the 'arrivals' and the radio channel as the 'server'. The analysis of a queuing system generally takes the form shown in Fig. 1. A model of the system is developed, based on characteristics such as the number of servers, behavior of the arrivals, the discipline by which the arrivals are selected for service, and the distributions which govern the arrival process and service time. When values are selected for whatever parameters are associated with the

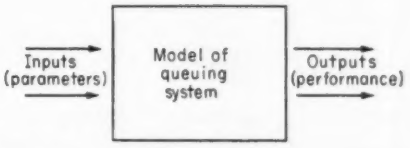


FIG. 1. *The process of analyzing a queuing system.*

model, characteristics of the operation can be found either analytically or by simulation. The most common parameters are the arrival and service rates. Among the characteristics measured are waiting times, service facility utilization, and the number of units waiting in the system. These and other characteristics would be considered as output from the queuing system model in Fig. 1, while the values of the parameters are input to the model.

In our attempt to model the queuing system of mobile radio communication we face several obstructions. In trying to specify the arrival process we find, as mentioned earlier, that we are unable to know every point in time when someone wants to send a message. If there has been no communication, then the next message sent does not have to wait, so we know that time of arrival. However, for any message which is sent immediately after another we do not know the time at which someone first wanted to send it—the time of its arrival to the queuing system. Another difficulty in the construction of the model is determining the discipline by which the next message is chosen for transmission. With multiple base stations and mobile units sharing the radio channel, at any point in time when communication is taking place we just never know how many messages are waiting to be sent or where they are. By what mechanism is it determined which message shall be sent? As we envisage a group of people with microphone in hand waiting for a chance to communicate, it is tempting to choose random selection for our model. However, anyone currently sending a message probably has a better chance of sending the next message if he has another one. This is due to the fact that other people will generally wait a few seconds to ascertain that the previous message has ended, while he knows he can just continue.

Other units, mobile or base, of the same company may share this same advantage. Thus, there may be a slight tendency to a priority discipline, with the priority going to whichever company is currently communicating. There is nothing formal or definite about the discipline of message selection, so whatever the model uses can never duplicate the actual operation.

Recalling that our purpose in analyzing the system is to assist in managing the communication channels used by mobile radio, we recognize that our interest lies in the output, not the model itself. In particular, we are interested in the delays encountered in sending messages. Since we cannot specify a model with all the components of the system, our analysis will proceed by finding a model whose output behaves similarly to that of the actual system.

THE ANALYSIS

Although our system is not completely structured, the radio channel is a single server. We begin our search for a model with similar behavior by finding a way to describe the arrival and service processes as best as we can.

Service time

Most of the messages were of fairly short duration; the average message length was 27 sec. A chi-square goodness-of-fit test was made for the hypothesis that message lengths were exponentially distributed as indicated by the data. The test indicated a close fit and leads to the use of exponential service time distribution in the model.

Arrival process

The actual arrival of messages to the queue to wait for transmission could not be measured directly since the arrivals occurred at all the company base stations as well as in the mobile units. The idle periods—the times with no communication—were measured. There is a problem of interpretation of these periods caused by the fact that someone waiting to send a message may listen for a few seconds to be sure he will not be interrupting a conversation in progress. On the other hand, it may be completely clear that the conversation has terminated, in which case he will immediately begin to communicate. In the dispatch offices both situations were observed.

The average of all idle periods was 43 sec. If the idle period is exponential, then the arrival process is Poisson. A chi-square goodness-of-fit test was made for the hypothesis, suggested by the data, that the idle times are consistent with a Poisson arrival process. For the data collected, the null hypothesis could not be rejected. Therefore it does not seem inappropriate to use a Poisson process in the model that is used.

While the data indicate a Poisson arrival process can be used, it is clear that the actual process cannot be observed. It is also clear that the process is not actually Poisson, for the preceding discussion described the way some of the brief idle times were observed while someone was actually waiting to send a message. Further, any company may have a greater message generation rate during the times it is already communicating knowing that it already has temporary control of the radio channel. As discussed elsewhere, when a radio channel becomes heavily congested, companies will refrain from sending relatively unimportant messages [6]. Thus, the arrival process will tend to be a function of the level of congestion. It is well to recall at this point that the concern is not whether or not the arrival process has been modeled exactly, since we cannot determine this, but whether or not the model selected produces output that accurately describes the actual output. The data on which the model will be based were collected for the channel operating at one particular level of congestion. Problems arising when using the model to evaluate changes on the radio channel will be discussed later.

The preceding section indicated that the message lengths observed could well be described as resulting from an exponential distribution. For a single channel queue with general independent input and exponential service, the times between successive departures are exponentially distributed if, and only if, the

arrival process is Poisson. Although earlier discussion provided some reasons why the unobservable message generation process may not have independent interarrival times, it is still possible that they could be independent. Were that the case, it would be useful to examine the times between successive departures from this queuing system.

The mean interarrival time was found to be 43 sec earlier in this section. It was decided that since the exponentially distributed times between successive message completions for a queue with exponential interarrival and service times would have the same parameter as the arrival process, the data would be tested for goodness-of-fit to an exponential distribution with a mean of 43 sec. The times between successive message completions actually had a mean of 49 sec, but the chi-square goodness-of-fit test still indicated that an exponential distribution with a mean of 43 sec could have resulted in the values obtained. This provides additional evidence that the arrival process may be Poisson, depending on the extent to which the assumption of independent interarrival times is appropriate.

The model

Earlier we discussed the fact that there is no strict queue discipline. Messages are selected neither completely at random nor by priority nor first come, first served; there is no formal queue. Let us see how well a model with the most basic discipline, first come, first served, describes two of the actual outputs for our queuing system, the busy periods and the waiting time. We will refer to this model as an M/M/1 queue, the standard queue notation.

Vol. 5
1977

The busy period in a queuing system is the time between idle periods. That is, when the system is empty a busy period begins with the next arrival and ends when the system is empty again. In our case a busy period is a time of continuous communication on the radio channel and certainly may consist of messages from several base stations and a number of mobile units. As mentioned earlier, a short period without communication could have occurred when there were no messages to be sent or when there was a message but someone was waiting to be sure the preceding message, sent by someone else, was completed. There was just no way to know which was the case. Since some of these brief periods of no communication may not have been legitimate idle periods, the real average busy period may be longer than calculated in this study.

For an M/M/1 queue with mean inter-arrival time $1/\lambda$ and mean service time, $1/\mu$, the mean of the busy period is readily obtained [3, p. 250]:

$$\text{average busy period} = \frac{1}{\mu-\lambda}$$

As discussed in the preceding sections, data observed on this radio channel indicate values for $1/\lambda$ and $1/\mu$ to be 43 sec and 27 sec, respectively. Using these

values in the expression above, the busy period of an M/M/1 queue would have an average of about 73 sec. In fact, this would be the average busy period for a single channel queue with Poisson arrivals and any independent and identically distributed service times with mean $1/\mu$. The busy periods which were observed during this study had an average of about 75 sec. Since earlier discussion showed there was little doubt concerning the message length distribution, it is appropriate to say that the busy period behavior of the system is that which would be expected from a system corresponding to the M/M/1 model.

Another queuing system characteristic that is generally of interest is the waiting time. In this study we are interested in the time that elapses from the moment someone wishes to send a message until the moment he begins to transmit the message. Here, as in many other queuing systems, the waiting time is a meaningful measure of system performance. A manager will often consider the system to be satisfactory if the delays are acceptable. In the case of our radio channel, lengthy delays result in complaints of radio congestion. As we will see in the next section, one reason that waiting time is a measure of performance is that there are usually costs associated with delay.

The density function of waiting time, $f(t)$, for an M/M/1 queue can readily be obtained [8]:

$$\text{Prob}\,(t = 0) = 1 - \lambda/\mu,$$
$$f(t) = (1 - \lambda/\mu)\lambda e^{-(\mu-\lambda)t} \quad t > 0.$$

The conditional density function for waiting time of messages that must wait is,

$$f(t|t > 0) = (\mu - \lambda)e^{-(\mu-\lambda)t}.$$

The expected values can be expressed in each case:

$$E(t) = \frac{\lambda}{\mu\,(\mu - \lambda)}.$$

$$E(t|t > 0) = \frac{1}{\mu - \lambda},$$

With the parameter values obtained from the data we find that for an M/M/1 queue, the expected wait is 46 sec, the expected wait for messages that must wait is 73 sec, and $f(t/t > 0) = 0{\cdot}0137 \exp(-0{\cdot}0137t)$.

The matter at hand is to determine whether or not these values describe the actual behavior of the system we are analyzing. Unfortunately we are unable to record actual waiting times and analyze them. We cannot know the moment at which any mobile unit or base operation desires to communicate, so when we hear a message being sent we generally do not know how long the sender was waiting. Of course we do know that any message following an idle period did not wait.

Unable to measure the waiting time behavior of our queuing system, an attempt was made to obtain some idea of what it may have been. The exponentially distributed inter arrival times represent a set of times until the next arrival from any point in time. Using these inter arrival times tape recordings which had been made of communication on the radio channel were started at random points in time. If the recording showed no communication taking place, a waiting time of zero was indicated. Otherwise, the recording was allowed to continue until the communication ended. This additional time was indicated as the waiting time. In this manner a sequence of simulated waiting times was constructed. The generation of messages was provided by actual data and the waiting times simulated by using actual tape recordings of communication on the radio channel. Of the 130 message arrivals simulated, 78, or 60 per cent, had to wait. With the data collected in this study we would expect that over the long run 63% of the messages would have to wait. The simulated waiting times for the 78 messages averaged 67 sec, close to the 73 sec a M/M/1 queue would average over the long run. The average wait for all 130 messages was 40 sec, close to the 46 sec an M/M/1 queue would average over the long run. A chi-square goodness-of-fit test indicated that the 78 waiting times could have come from the distribution which results from an M/M/1 queue. For an exponential distribution, the standard deviation is equal to the mean. With the sample of simulated messages which waited, the S.D. was 63 sec, close to the mean of 67 sec.

Earlier discussion indicated that the actual mechanism by which the next message was chosen for transmission was not definite. With messages generated at a number of mobile units and base stations, random selection might seem appropriate. The average waiting time would be 46 sec with random selection, just as with the first come, first served mechanism. The latter case would result in a standard deviation of waiting time of 68 sec. The standard deviation of waiting time with a random service mechanism can be determined from the first and second moments [1, p. 236]. The model with random service would have a waiting time S.D. of 88 sec. The 130 simulated messages had a waiting time S.D. of 59 sec. This 'actual' system performance is closer to that indicative of a first come, first served discipline than that indicative of random service.

USING THE MODEL

In the previous section we found that the model of an M/M/1 queue gave a good description of some of the measures of behavior of the radio channel. Now we can use that model to determine the potential effect of changes. This will be done by examining the delay encountered in sending a message.

When there is delay in sending a message, there are different possible results.

For example, the delay may cause a truck to go farther in the wrong direction, costing driver time and vehicle mileage. At other times the delay may require the driver to wait idle until the message can be sent. The detailed analysis elsewhere presents models of the various possibilities [5]. For simple illustration consider here a company sending 80 messages per day. Suppose that for 40 of these messages, a delay will result in idle driver time, having the effect of lengthening his working day. If \$9 is the hourly overtime cost of a driver and W_q is the average delay (in hours) in sending a message, then the average daily delay cost for such messages is (40) (\$9) (W_q). Earlier we evaluated W_q as 46 sec. The average daily delay cost for this company would be (40) (\$9) (46/3600) = \$4.60.

Suppose in the future a thirteenth company is assigned to the radio channel. Later, a fourteenth might be added. We can use the model to evaluate in advance the effect of these new companies. After such an evaluation it might be deemed inappropriate to add the new companies. The effect on delay depends on the increase in messages, not the number of additional companies. Assuming the new messages average 27 sec in length, as did the messages of the 12 current companies, the average daily delay cost for the one company we have illustrated would increase as shown in Fig. 2. In using Fig 2. it should be recalled that under heavy congestion the companies may change their operations by eliminating some messages. When the thirteenth company is added to the channel, the other companies may reduce their message generation rates in an effort to lessen the effect of the additional company. The delay cost in Fig. 2 would not reflect this reduction. On the other hand, by eliminating some messages some loss of the radio's value is incurred. Figure 2 should provide a general guide in attempting to understand the effect of increasing the communication on a radio channel.

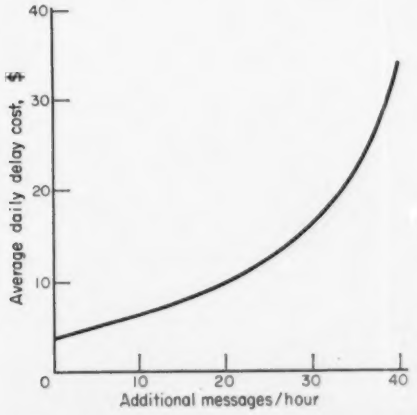


FIG. 2. *The effect of additional communication on one company's daily cost of delay.*

SUMMARY

A mobile radio channel shared by twelve companies was investigated. It was observed to be a queuing system for which some components and characteristics could not be specified or measured. A general procedure for analysis under such circumstances—the use of a model whose output was similar to the actual system—was presented and was illustrated for this case. The use of such a model to evaluate potential changes in the complex, partially unobservable queuing system was demonstrated.

REFERENCES

1. COOPER RB (1972) *Introduction to Queueing Theory*. Macmillan, New York.
2. GRAFF G (1961) Simple queuing theory saves unnecessary equipment. *Ind. Engng*. **3**(2), 15–18.
3. GROSS D and HARRIS CM (1974) *Fundamentals of Queueing Theory*. John Wiley, New York.
4. HILLIER FS (1963) Economic models for industrial waiting line problems. *Mgmt Sci.* **10**(1), 119–130.
5. HILLIER FS (1965) Cost models for the application of priority waiting line theory to industrial problems. *J. Ind. Engng* XVI(XVI, No. 3, 178–185.
6. PLOTKIN HA (1973) Mobile radio—a pickup and delivery tool. *Logistics Transportation Rev.* **9**(4), 357–366.
7. SAATY TL (1961) *Elements of Queuing Theory*. McGraw-Hill, New York.
8. SASIENI M, YASPAN A and BLOCK L (1959) *Operations Research*. John Wiley, New York.

ADDRESS FOR CORRESPONDENCE: *Howard A Plotkin, Assistant Professor, Quantitative Management Science Department, University of Houston, Houston, Texas* 77004, *USA.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

Memoranda

The Optimum Planning Period Problem

THE OPTIMUM planning period problem arises when service capacity must be expanded to cater for an indefinitely growing demand. Each expansion caters for future as well as current demand, thus considerable idle capacity may be incurred. As demand continues to grow the proportion of service capacity idle decreases until a further expansion becomes necessary. The object is to find the capacity expansion programme which minimises long term capital costs. In practice only the optimum size of the first capacity expansion is required; subsequent further optimum expansions are estimated later in light of potentially changed conditions.

Past work on this problem comes largely from communications engineers interested in the optimal expansion of telephone networks and other types of plant. The solution when demand grows linearly has long been known [1]. A calculus technique has been used to obtain a good approximation to the exponential demand case [2], and a number of alternative specialized forms have also been solved [3]. However, the following dynamic programming formulation may be applied to more general and realistic demand and supply situations.

The problem is solved in two stages. First it is solved with a finite horizon, and then the response of the solution to a lengthening of the horizon period is estimated.

The horizon period is divided into a number of equal intervals, and a decision is made at discrete time intervals whether or not to make a capacity expansion to satisfy demand for some number of subsequent intervals. The length of each interval is chosen in accordance with accuracy requirements, budgeting practices, and so forth.

We will consider a possible expansion at time period $t = r$. Let us define $D(t)$ = demand for service (capacity requirements) at time t_j; A = variable cost per unit capacity; B = fixed cost incurred at each capacity expansion; n = number of fixed periods into which the finite horizon is divided; w = discount rate per period; r = parameter referring to the number of time periods elapsed from $t = 0$, i.e. $0 \leqslant r \leqslant n$; s = number of periods beyond $t = r$ for which capacity exists at $t = r$ (without expansion at $t = r$), i.e. $0 \leqslant s \leqslant n - r$; x_r = number of periods beyond $t = r$ for which capacity is expanded at $t = r$ (when $s = 0$), i.e. $0 \leqslant x_r \leqslant n - r$; $C_r(s)$ = cost of following an optimum expansion programme from $t = r$ to $t = n$ when there exists capacity s at $t = r$. Now, two alternative cases are identifiable.

1. If capacity exists for at least one period ($s \geqslant 1$), then no expansion is necessary ($x_r = 0$). The discount factor implies that present value costs are minimised when expansions are delayed as long as possible and hence the optimal expansion at $t = r$ with capacity $s \geqslant 1$ is no expansion. The optimal cost at $t = r$ with capacity $s \geqslant 1$ is the discounted optimal cost at $t = r + 1$ with capacity $s - 1$.
2. If at $t = r$ there is no capacity for further periods ($s = 0$) then expansion is essential ($x_r \geqslant 1$). To satisfy demand from $t = r$ to $t = r + x_r$ capacity is increased by $D(r + x_r) - D(r)$ at a cost of $A[D(r + x_r) - D(r)] + B$. The optimal expansion at $t = r$ is that which minimises the expansion costs for x_r periods plus the discounted optimal costs at $t = r + 1$ with capacity $x_r - 1$. This optimisation may be expressed as a *DP*, i.e.

Find x_0 which minimises $C_0(0)$ if $C_n(0) = 0$ and for $r = 0$ to $r = n - 1$.

(a) $n - r \geqslant s \geqslant 1, x_r = 0,$
$C_r(s) = w\,C_{r+1}(s - 1);$

(b) $s = 0, n - r \geqslant x_r \geqslant 1,$
$C_r(s) = \underset{x_r}{\text{MINIMUM}}\{A[D(r + x_r) - D(r)] + B + wC_{r+1}(x_r - 1)\}.$

As the finite horizon is increased from a low of zero the length of the first planning period, x_0, increases linearly until the horizon is sufficiently large to enable the discount factor to become effective. Once this has occurred x_0 no longer increases but falls slightly, and then remains constant for a

wide range of the horizon. This constant value can be taken as the solution to the infinite horizon problem.

The above dynamic programming formulation has been applied to number of different demand functions. The formulation was found to be easily adapted to alternative conditions and provided rapidly converging results under all combinations of

(a) linear, exponential, saturating, or discontinuous demand,
(b) continuous or discrete capacity increments, and
(c) continuous or discrete decision or budgeting periods.

The results in the exponential and linear cases appear to be similar to those given graphically by McDowell [2] and Rapp [1], respectively.

In the most common case of exponential demand the above formulation avoids to very real computational and discontinuity difficulties entailed in McDowell's calculus approximation technique. Further it was found that realistic solutions could be obtained even under conditions of exponential demand with a growth rate greater than the discount rate. Rapp pointed out that under these conditions the cost of future expansions increase with time, and hence the cost of expansions taking place far into the future will dominate the total cost and correspondingly bias the first planning period. The present work has confirmed these anticipations, but found this effect to occur only when the horizon period is very large indeed. x_0 is constant for a large range of the horizon < 150 years irrespective of the relative values of demand growth rate and discounting rate; when demand growth rate > discounting rate then x_0 rises very rapidly above the previously constant value for an horizon much larger than 150 years. The solution to the long term planning period problem can certainly be taken as that constant value of the first planning period which occurs for a wide range of the horizon.

REFERENCES

1. Rapp Y (1961) Extensions of telephone plant with regard to the value of the subscriber's time, Part II, *Ericsson Tech.* **17**(1), 62.
2. McDowell I (1960) The economical planning period for engineering works. *Ops Res.* **8**(7), 533.
3. Kurahashi K (1964) The optimal expansion of production facilities under limited capital. *Rev. Elect. Commun. Lab, Tokyo* **12**(3–4), 151.

PE Becker
(*August* 1976)

Centre for Economic Research
Swiss Federal Institute of Technology, Zurich
ETH-Zentrum
8092 *Zurich*
Switzerland

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

Could OR be a Science?

IN AN OMEGA editorial in 1975, Eilon [5] considered the question "How Scientific is Operational Research?", and thereby provoked much 'Feedback' in subsequent editions of the journal [6]. It is not our intention to reopen that debate; we wish rather, to examine an important issue underlying it. Our question is whether OR is, or ever could be, a science in the same sense as we understand Physics, Chemistry or Biology? We raise this question because we suspect that within the OR community, the use of some scientific procedures is widely assumed to give OR the status of a science.

We do not accept this assumption. Therefore, our initial aim is to show that, despite the undisputed use of some scientific procedures, OR is not a science. More importantly, we wish to demonstrate the serious consequences which follow from the presumption that OR is a science.

What are the characteristic features of the mature sciences? For present purposes we will consider the following list.

i. A clear focus on a major sub-set of phenomena of human concern.
ii. A working set of paradigms (such as atomic and molecular theory in Chemistry), and an associated language, which are sufficient for description, conjecture and debate over the whole range of phenomena.

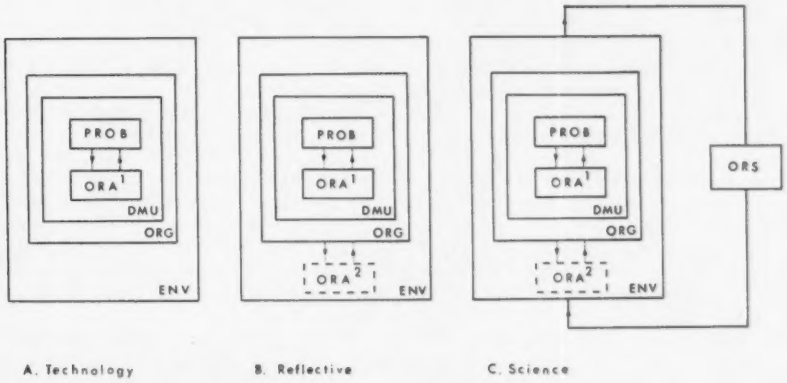


FIG. 1. *OR orientations. In this conceptual box diagram descriptive system (see Bowen and Harris*[3] *for details), it is intended to depict the various subsystems of interest in relation to each other* (ENV, *environments*; ORG, *organisations*; DMU, *decision-making units*; PROB, *perceived problems*; ORA[1], *operational research analysts*; ORA[2], *the same but as a conceptual reflection whilst still basically concerned with* ORA[1] *issues*; ORS, *operational research scientists deliberately attempting to view decision-processes from outside). The arrowed lines indicate flows of information/action relevant to the argument presented here.*

iii. A widely-accepted system of testing rival theories.

iv. The overall aim of producing general causal explanations of the mechanisms underlying the patterns which arise in the area of concern (that is, descriptions of how parts of the system work, which in total provide a coherent framework of understanding).

If we examine the activities of OR analysts (ORA^1 in Fig. 1A) we see that their endeavours do not have these features.

i. The subject of study is merely certain limited types of decision-problems [8] perceived by a restricted group of decision-makers.

ii. There is no working set of paradigms and language sufficient for general description, conjecture and debate even across this limited range of subject matter.

iii. Rigorous theory-testing is infrequently considered necessary.

iv. The overall aim is *not* to produce general causal explanations of how the system works. Rather, it is to contrive devices by which, in relation to the individual decision-problem of the moment, the system can, in some sense, be supposed to work 'better'.

Surely, OR is a technology equivalent in many ways to the various branches of engineering? OR analysts confront decision-problems in the real world and deal with them in a strictly pragmatic manner. The most important feature which we would consider a fundamental criterion for a science, is iv. The concentration of OR on improvement, in relation to the individual decision-problem of the moment, contrasts with a true science where the search for general understanding is central. In short, OR, like other professional advisory disciplines but unlike a science, is fixated on the decision-problem.

We find nothing disturbing in OR being a technology. The world needs engineers as much as it needs physicists. The difficulties only arise if we imagine that we are practising a science when it is, in fact, a technology. Obviously, if we imagine that we have a science, there will be no apparent need for the development of one, but if we have no science, we are condemned to have a poor technology (consider the plight of engineering without physics).

The point is not lost, at least on some of the OR community. White [10], in his recent book "*Decision Methodology*", has shown that even for simple problems the analyst may face almost impossible tasks in deciding, for example, how to decide what the problem is, or how to go about selecting the appropriate technique to deal with it. We would also point, as evidence of the awareness of major difficulties with present practices, to the widespread concern over "behavioural problems" [3].

The question that follows is whether we are condemned to this situation by factors beyond our control, or by our own choice. We would answer that it is entirely by our own choice. The history of many technologies shows that they arose as practical endeavours and only later were codified as sciences, when the possibilities of more systematic and rapid progress were grasped. Nevertheless, we think that whatever the OR analyst does in an attempt to distance himself conceptually from his present individual decision-problem (ORA^2 in Fig. 1B), there is little chance of the rapid build-up of systematic, formulated knowledge whilst we all remain within the present technological orientation. One step forward would be for some members of the OR community to place themselves in the position of the scientist further along the ORA^1, ORA^2, ORA^n series. From that position it might be possible to discover more effectively how the ORA^1 analyst operates. However, this seems to us to be insufficient as an attempt at a true science. We do not think that the analyst's role can be examined properly outside the context of the whole decision-process. We feel that the best development would be for the OR community to assist in the establishment of a science focussed on the analysis of decision-processes (ORS in Fig. 1C). Present day OR might then grow within an associated, broader technology concerned with the design of decision-systems.

We are not suggesting the introduction of some new dispassionate, 'objective super-analysts' [4]. The issue is whether we should continue to be satisfied solely with our present disjointed attempts to make a limited set of individual decision-problems better, or if we should also attempt to construct an adequate science and technology. With such a science and technology we could then make a much greater impact on the real world problems which concern us.

Considering again the characteristics of

Vol. 5
1977

mature sciences, we suggest that OR scientists might be able to construct a science with the following features.

i. Decision-making processes in organisations as the major sub-set of phenomena of concern.
ii. Systems theories [1] as the basis of the paradigms and language of description, conjecture and debate. These theories are well-developed, but infrequently associated with the third feature, a rigorous system of theory testing.
iii. Theory testing dependent on the methods of strong inference [7], linked to the use of some forms of controlled gaming [3, 9]. Gaming is justifiably not well-regarded at present, but with care our use of controlled gaming could parallel the standard practice, in the mature sciences, of using appropriate laboratory systems for rigorous theory testing.
iv. The aim of producing general causal explanations of how decision-processes work. These would then form an adequate framework of understanding for the associated technology of decision-system design.

These suggestions may sound far-fetched; but without such developments OR must remain subject to chance and luck for its future progress. If, however, we can reach a consensus that OR is a primitive technology in need of a science, then the programme of development required is straightforward.

i. Further work as necessary on the systems paradigm and language.
ii. Intensive study of the construction and use of controlled games as experimental environments for theory testing.
iii. Examination of relevant sources for regularities in areas of interest. Again there is nothing exceptional in this suggested search for patterns or 'primitive images' [9], but we should perhaps do well to recognise the prescientific state we are in.
iv. Use of the perceived patterns as the basis for the formulation of testable theories in the standard paradigm and language.
v. Testing of the rival theories proposed to explain the perceived patterns.
vi. Attempts to link i. to v. with the important real world problems which generated the search for the patterns in the first place.

This will be a monumental task as we are proposing a new social science [2], and we are suggesting a programme which no social science has ever accomplished. Yet eventually it will have to be done.

Our thesis and proposals will be badly received in some quarters, but it should at least be noted that we are not attempting to denigrate the activities of technologists. We are simply suggesting that they, and the people affected by their activities, may be able to make more satisfactory decisions (by their own criteria) if they are able to operate within the context of the general framework of understanding provided by a true science. Moreover, our message is not primarily directed at the real world practitioners of OR, but at some of their academic colleagues. Certainly, a primitive technology has to teach the disjointed collection of techniques which are sometimes required by the students when they go to work now; but must our universities restrict themselves quite so rigidly to this simple task when OR analysts are encountering (and will increasingly encounter) problems which cannot be tackled by such methods? Surely, in our university OR departments there is a great need for the development of both teaching and research orientated towards the sort of science and technology proposed here?

ACKNOWLEDGEMENTS

We would like to thank several friends—A Berresford, KC Bowen, R Faulkner, JI Harris, KJ Radford, BHP Rivett and HRW Watkins (although their views do not necessarily agree with our own), for their comments on earlier drafts of this script.

REFERENCES

1. Ashby RW (1964) *An Introduction to Cybernetics*. Chapman & Hall, London.
2. Bevan RG (1976) The Language of Operational Research. *Opl Res. Q.* **27**(2), i, 305–313.
3. Bowen KC and Harris JI (1976) Problems with People; Decision, Conflict, Language and Measurement. Paper to be presented at NATO Conference on "Environmental Assessment of Socio-Economic Systems", Istanbul, October, 1976.

4. EASLEA B (1973) *Liberation and the Aims of Science*. Chatto & Windus for Sussex UP, London.
5. EILON S (1975) How scientific is OR? *Omega* **3**(1), 1–8.
6. FEEDBACK (1975–76) *Omega* **3**(5), 519–23; **4**(1), 9–19.
7. PLATT JR (1964) Strong inference. *Science* **146** (3642), 347–353.
8. RADFORD KJ (1975) *Managerial Decision-Making*. Reston, Virginia.
9. TEDESCHI JT, SCHLENKER BR and BONOMA TV (1973) *Conflict, Power and Games*. Chapter 7, Games as Research Tools. Aldine, Chicago.
10. WHITE DJ (1975) *Decision Methodology*. Wiley, London.

MR Dando
A Defrenne
RG Sharp
(*September 1976*)

Department of Operational Research
University of Sussex
Falmer
Brighton BN1 9RF
UK

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

Ex ante–Ex post Eurobond Experience: Some Canadian Evidence[1]

PRIVATE firms with assets and income denominated in more than one currency have long followed various hedging strategies, including offshore borrowing, in an effort to control their exchange risk exposure. Sovereign states often borrow abroad for balance of payments or monetary policy reasons. Political subdivisions of such states, however, are generally thought to expose themselves to the exchange risk inherent in a foreign loan either because the cost of funds in the domestic market is perceived as too high and/or the absorptive capacity is too low. Furthermore, it should be stressed that the true cost of a foreign issue reflects not only the coupon rate but also issuance costs and exchange risk. With regard to the last, basic portfolio theory indicates that the hierarchy of desirability for denomination would be (1) one's own currency; (2) a closely related currency; (3) an uncorrelated currency; and, finally, (4) an inversely correlated currency. We should expect to find issuance costs related to the size and competitiveness of the market of issue. Finally, we should expect the coupon rate to reflect interest rate conditions in the market of issue and any special attributes of the particular issue.

The implied trade-off between coupon rate, issue cost, and exchange risk can become even more subtle. Swoboda [10, p. 18] has defined a currency which is widely held and used in international transactions as a vehicle currency; for the post-war period, the US dollar is the classic example. He went on to suggest that savings in transactions costs allowed lower rates to be charged upon liabilities denominated in vehicle currencies and that this reduction constituted a denomination rent to residents (and seigniorage to the government) of vehicle currency countries. Finally, he implied that others who were prepared to endure the exchange risk of vehicle currency denomination could bid away some of these rents.

The Eurobond market provides a unique combination of the traditional foreign loan with some aspects of the use of a vehicle currency [3, 5]. In sum, a prospective borrower (who might or might not have been forced out of his domestic market for reasons of absorptive capacity anyway) confronts not only the exchange risk of a traditional foreign loan but also the possibility of reaping denomination rents in the form of lower coupon payments by the use of a more 'popular' international currency. If such rents do exist, they would provide at least partial compensation for the exchange risk borne; indeed, if they were great enough, they could induce a borrower out of his own domestic market.

The empirical determination regarding whether a given borrower is forced out of his domestic market or borrows abroad voluntarily is most difficult and not addressed here. Given a decision to borrow abroad, however, we can investigate the question of whether nominal cost patterns conform to the more traditional exchange risk–portfolio theory model or whether the existence of other factors (e.g. the above-mentioned rents) distorts this relationship. We can also investigate similar questions in regard to lenders.

Canadian provincial authorities borrow a great deal of money, and the Canadian capital market is not very broad. Hence, at least some of their extensive foreign borrowings could be attributed to the absorptive capacity of the domestic market. Based upon several case studies over the 1969–74 period, we shall evaluate the *ex ante* and *ex post* cost

[1] The authors are Professor of Finance, University of Southern California and Research Analyst, Royal Bank of Canada, Montreal, Quebec respectively. This work was completed while both were associated with McGill University, the use of the facilities of which is acknowledged. An earlier version of this paper was presented before the Academy of International Business, Dallas, December 1975.

of these borrowings and compare the rankings to the alternatives outlined above. Next, we shall also evaluate the position of some possible purchasers of these bonds. Finally, some policy implications are drawn.

Because a bond is generally issued for an original maturity of 15–25 years and the world can change a great deal in the interim, evaluations of lender and borrower experience must inevitably deal with the fact that the outcome will never correspond to the expectation. The short history of post-World War II international borrowing precludes us from observing even one complete cycle of borrowing and repayment. We chose a 5 year horizon as being long enough to include some unexpected events and yet short enough so that the decision point on flotation was still reasonably current. Hence this section evaluates the options facing the provinces during the summer of 1969 from the perspective of early 1974.

The domestic Canadian market provides a useful benchmark for this comparison, as obviously no exchange risk would be incurred upon borrowings there. During 1969, yields in the Canadian market rose consistently and for provincial issues, averaged about 8·35% during the summer [8]. Abstracting from questions of absorptive capacity and including normal flotation expenses, we estimate a corresponding cost of funds to provincial borrowers of 8·7%.

Some amplification of this cost computation might be appropriate, as we use it throughout this paper. The cost to borrowers is defined as the internal rate of return which equates the proceeds of the issue, net of flotation expenses and expressed in Canadian dollars, with the Canadian dollar equivalent of the scheduled interest and sinking fund payments. The impact of changing exchange rates upon the latter provides the *ex post* cost estimates we report below. Finally, the same basic computation employing the lender's currency and the gross issue price provides the basis for the lender return computations in the next section.

Thus, from the above, we conclude that a provincial borrower at the margin could have avoided exchange risk altogether at an annual effective cost of about 8·7%. The next least risky alternative would appear to have been US dollar denomination, especially in view of the high degree of correlation historically between the US and Canadian economies and currencies. In this case, borrowers would have had a choice between offering a dollar Eurobond or a foreign bond in New York; in the summer of 1969, however, rates were quite similar between the two and we shall only evaluate the latter [7, p. 121]. The New York option was popular, as Canada has consistently offered $800–1000 million per year of bonds in the US over the last decade (and double this amount in 1975 [2]), with upwards of half offered by provinces and half bought by US residents [cf. 9]. The salient features of a Hydro Quebec issue are shown in Table 1, with the conclusion that a weakening of the US dollar over the 1969–74 period actually caused a reduction in cost to the borrower. Because, however, this finding is quite sensitive to the date chosen for the evaluation, the safer conclusion is that the opportunities for either gain or loss to the borrower from parity changes were not very great.

Somewhat greater risk exposure would be provided by a currency essentially uncor-

Vol. 5
1977

TABLE 1. COST TO BORROWERS

Issuer Value	Hydro Quebec U.S. $50 m	Manitoba EUA 15 m (Can $16.2 m)	Newfoundland DM 80 m (Can $21.6 m)
Coupon	8·75%	7%	7·25%
Issue price	100	97	97·5
Life (years)	30	20	15
Issued	7/1/69	6/17/69	7/31/69
Cost to maturity: at issue	9·2%	7·85%	8·3%
At 1/1/74	8·4%	9·2%	12%

related with the Canadian dollar; we have chosen the unit of account (EUA). A unit is equal to the 'old' US dollar gold parity of the currencies of the 17 members of the European Payments Union. To avoid proscription as a gold clause, EUA bonds provide that, if all 17 currencies change their gold parities and two-thirds change in the same direction, then the unit shall be adjusted in the direction of the majority by the amount of that currency which changed the least. For this computation, currencies which cease to have a gold parity are dropped and, if all cease, the unit is defined in terms of the most stable currency. The EUA has thus been lauded as minimizing exchange risk for both lenders and borrowers [6]; in opposition it has been noted that the formula is quite confusing and, perhaps as a result, has not caused any significant reduction in coupon for borrowers [1, p. 219]. The features of a Manitoba EUA issue are shown in Table 1. The increase in borrower cost over the evaluation period is attributable to the fact that the EUA appreciated twice against the US dollar (i.e. the dollar devalued against gold while the EUA did not) to the extent of 20·6%, while the Canadian dollar had recovered only half of this amount against the US dollar by January, 1974.

Finally, the greatest exchange risk exposure would be incurred by denominating liabilities in a currency inversely correlated with the Canadian dollar. The currencies with the greatest revaluation potential in 1969 included the deutschemark (DM) and Swiss franc [5]. On the other hand, of course, both the availability of funds and coupon rates reflected the popularity of these denominations with lenders. The Swiss government controlled access and use of the franc, but a booming Eurobond market existed in DM. The features of an issue by Newfoundland are depicted in Table 1, along with the extensive revaluation losses suffered between 1969 and 1974.

Surveying the cost revisions in Table 1, we find that the only *ex post* deviation from the exchange risk–portfolio model implications is represented by New York and is also rather small. Although generalizations from case studies must be undertaken and interpreted with care, they are presumably no worse than those drawn from no empirical evidence whatever. Denomination rents might explain the slight deviation favoring New York above; however, the more traditional considerations of market breadth and competitiveness seem adequate as explanations.[2] Indeed, the traditional exchange risk model appears well supported by the evidence. It would follow from our analysis that borrowers seeking foreign loans should not anticipate any significant rents to offset the exchange risk they must bear. Furthermore, the above evidence with respect to borrowers is consistent with at least two elements of the conventional wisdom [4, 5]:

1. The coupon reduction on strong-currency (i.e. DM) Eurobonds does not fully compensate for the appreciation potential of the currency of denomination.
2. The EUA does not appear to offer adequate coupon reductions to borrowers to compensate for the depreciation potential foregone.

In addition to the data above, we also computed the *ex post* return to lenders for the two Eurobond issues. In the case of both issues, it was assumed that parities prevailing as of early 1974 would hold (or, more generally, would be unbiased estimators of future parities) for the remainder of the issues' lives. In the case of the EUA issue, we also assumed that payment would be demanded in that currency relatively most valuable (which we further assumed to be the DM in 1974–81). Thus, we obtain the yields to average life, in terms of lender currency, based upon initial cost revised for 4½ years' experience shown in Table 2.

From the lender perspective, it will be noted that holders of weaker currencies received a higher return than holders of strong currencies, although their opportunity costs (i.e. domestic interest rates) could also be expected to be higher. Absolutely, the DM issue was more desirable for lenders than the EUA issue; not only does this suggest that No. 1 above applies symmetrically to borrowers and lenders, but it also lends support to a third element of the conventional wisdom [cf. 4, 5]:

3. The EUA bonds are especially undesirable for strong-currency lenders. In this regard, it is instructive to note that the yield on the EUA bonds to Swiss, Dutch, and, especially, German holders declined,

[2] This slight differential might also result from the exemption of interest payments to US investors by provincial and municipal authorities from the 15% Canadian withholding tax [2].

TABLE 2. REVISED YIELD TO LENDERS

	Manitoba EUA	Newfoundland DM
Belgium	7·25%	9·8%
France	9·0	12·5
Germany	5·0	7·7
Italy	9·1	12·5
Netherlands	6·7	9·2
Switzerland	6·1	8·4
United Kingdom	9·7	13·0

on a historical cost basis, from 1969 to 1974. Hence, in any sort of "lenders' market," the lack of popularity of EUA bonds is not difficult to understand.

Finally, the case studies presented here also illustrate an interesting theoretical point relating to offshore borrowing by political subdivisions of a sovereign state. If these bonds do indeed stand on their own, then even the *ex post* results may be rationalized in terms of equi-marginal borrowing terms and market equilibrium. It need merely be assumed that the various provinces could not obtain additional funds in the domestic market at our estimated 8·7% and were forced, at the margin, into higher cost markets consistent with their perceived risk. The latter would constitute a premium paid for this risk, however, and involve a social cost.

If, however, the Government of Canada is (or is perceived to be) the *de facto* ultimate guarantor of these bonds, then our sample constitutes a homogeneous risk class, no premium should have been forthcoming, and equi-marginal conditions were violated (at least *ex post*). Formal efforts to eliminate this premium, however, would require the centralization of offshore borrowing and, quite likely, a federal guarantee. The latter would, in turn, require either strict federal allocation of proceeds or else an unlimited guarantee on provincial borrowings; either alternative would create severe political problems.

If the two alternatives above were exhaustive, then it might be concluded that the risk premia constituted a second-best solution to the restriction of provincial borrowing. There remains the possibility, however, that the market charges premia and still views Canada as an implicit guarantor. From the standpoint of the federal government, this constitutes the worst of both worlds; a good empirical test, unfortunately, would require a provincial default and subsequent market reaction (as in the recent New York City–New York State case). In any event, no sovereign state can afford to ignore the implications of offshore borrowing by its subdivisions.

Vol. 5
1977

REFERENCES

1. ALTMAN O (1965) The integration of European capital markets. *J. Fin.* 209–21.
2. Canadian borrowing in US markets is running at more than double 74 pace. *Wall Street Journal* (October 5, 1975), p. 13.
3. DUFEY G (1969) *The Eurobond Market—Function and Future.* Gr. School of Business Administration, University of Washington, Seattle.
4. FINDLAY MC (1974) Considerations for the multinational firm seeking funds in the Euro-capital markets in *Management of the Multinationals* (Eds. SETHI and HOLTON) pp. 292–312. Free Press, New York.
5. FINDLAY MC (1970) The evolution, state, and implications of the Eurobond Market. Unpublished dissertation, University of Texas.
6. *The Mechanics of Eurobond Issues* (1967) Banque de Bruxelles, Brussels.
7. MENDELSON M (1972) The Eurobond and capital market integration. *J. Fin.* 110–126.
8. *Statistical Summary*, Various Issues. Bank of Canada, Ottawa.

9. *Survey of Current Business*, Various Issues. Department of Commerce, Washington.
10. SWOBODA AK (1968) *The Euro-dollar Market: An Interpretation.* International Finance Sec., Department of Economics, Princeton University, Princeton.

M Chapman Findlay III
ER Thompson
(*September 1976*)

Department of Finance
University of Southern California
School of Business
Administration
University Park
Los Angeles, CA 90007
USA

l. 5
77

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 1, 1977. Pergamon Press. Printed in Great Britain

EOQ Model with Continuous Compounding

GENERALLY the average or expected cost has been used in analyzing inventory systems. In doing so, insufficient consideration is given to the time value of the money. The opportunity cost which is claimed to incorporate the time value of the money is often vaguely defined and is based on annual compounding. In most inventory systems transactions occur throughout the year; these circumstances suggest a continuous flow of money, for which continuous compounding and discounting is more realistic. Hadley and Whitin [2] discuss the possibility of using present worth approach, but suggest that the results will not be significantly different from those obtained using average cost method for all practical purposes. This note presents a procedure to calculate the present worth of the cost of operating an inventory system, and compares the cost with that of the average cost method.

The proposed approach uses two continuous rates, one for opportunity cost and the other for real-out-of-pocket inventory costs; these two rates will be called investment rate (r) and inventory rate (e) respectively. The two separate rates are needed in order to assess the present worth of the money including the real-out-of-pocket costs of maintaining an inventory. For mathematical convenience the inventory rate will be written as a multiple of the investment rate.

DEVELOPMENT

The development utilizes the usual assumptions associated with the economic-lot-size model [2]. In addition, it is assumed that the sales price of a unit is constant and the use of money is continuous. Further notation used in this paper is as follows:

λ = demand/year; Q = reorder quantity; A = cost of placing an order; c = cost/unit; s = sales price/unit; $k = \rho/r$; i = inventory carrying rate [$i \equiv r(1+k)$]; I = annual inventory carrying rate; P_τ = present worth of the total profit/year and k_τ = present worth of the total cost/year. Since I, in the conventional approach [2], incorporates both real out of pocket costs and opportunity costs, it can be thought of as the equivalent annual rate corresponding to i, i.e. $I \equiv e^i - 1$.

The present worth of the profit during the first cycle is

$$p_1 = \int_0^{Q/\lambda} \lambda s\, e^{-rt} dt - (Qc + A) - \int_0^{} ckr\,(Q - \lambda t) e^{-rt} dt$$

or

$$p_1 = \frac{\lambda s}{r}(1-e^{-rQ/\lambda}) - Qc(1+k) - A + \frac{k\lambda c}{r}(1-e^{-rQ/\lambda}).$$

As there are λ/Q cycles in a period, the present worth of the total profit during a period is

$$P_\tau = p_1 \frac{1-e^{-r}}{1-e^{-rQ/\lambda}} = \left[\frac{\lambda s}{r} - \frac{Qc(1+k)+A}{1-e^{-rQ/\lambda}} + \frac{k\lambda c}{r}\right](1-e^{-r}). \quad (1)$$

Since the revenue term in P_τ is independent of Q, minimizing P^τ will also minimize the cost, K_τ, i.e.

$$K_\tau = \left[\frac{Qc(1+k)+A}{1e^{-rQ/\lambda}} - \frac{k\lambda c}{r}\right](1-e^{-r}). \quad (2)$$

Differentiating P_τ with respect to Q and equating the result to zero gives

$$e^{rQ/\lambda} - \left(1 + \frac{Qr}{\lambda} + \frac{Ar}{\lambda c(1+k)}\right) = 0. \quad (3)$$

An exact analytical solution of (3) for Q is not possible but a numerical solution Q_e may be obtained using the Newton–Raphson method. However, an approximate solution (within about 2%) can be obtained by using a second order approximation for the exponential term, i.e.

$$Q_w \simeq \sqrt{\left(\frac{2\lambda A}{r(1+k)c}\right)}. \quad (4)$$

One may also assess the value of Q using the conventional model, Q_c, i.e.

$$Q_c = \sqrt{\left(\frac{2\lambda A}{Ic}\right)}. \qquad (5)$$

It can be seen from equations (4) and (5) that the numerical values of Q_w and Q_c will not be significantly different if the value of I is selected such that it is equivalent to the sum of continuous rates r and ρ. A number of problems were solved and the variation of Q_w from Q_c ranged from 0 to 7%. Furthermore, the variation of the present cost for changes in Q was investigated and found to be very small in the neighborhood of the optimum Q. This suggests that any one of two reorder quantities, Q_w or Q_c, can be used as the reorder quantity without any significant influence on the cost.

DISCUSSION

In the light of obtaining more or less the same optimum reorder quantity using the present worth approach, a basic question arises—what good is this approach? The answer to this question may come from understanding the nature of these two costs. The present worth cost is value of the money now, which is equivalent to all expenditures over a period; it takes into consideration the time of occurrence of each expenditure. This is clearly not true for the average cost and hence the present worth cost is a closer representation of the real cost situation. To get a numerical comparison of the two costs, an example problem was selected from [1]. The data are $A = \$150$, $c = \$20$, $\lambda = 508$, $I = 20\%$. I represents conventional inventory carrying rate, which incorporates both opportunity cost and out-of-pocket inventory costs. In the absence of any real breakdown of I into its two main components, three cases were considered: first, where out-of-pocket costs were zero (i e. $k = 0$); second, where out-of-pocket costs were one quarter of opportunity cost ($k = 0{\cdot}25$); third, where out-of-pocket costs were equal to opportunity cost ($k = 1$). In order to make the results for both approaches equivalent, a corresponding r was calculated such that, for each k, $I = \exp[r(1+k)] - 1$. The values of r thus obtained are 18·23%, 14·58% and 9·11% respectively.

The optimum reorder quantities, $Q_w = 203$ and $Q_c = 196$ units (from (4) and (5)). The average cost/year is equal to \$11,046; the present worth costs/year for the three cases are \$10,083, \$10,252, and \$10,527. These values of present worth cost are lower than the average cost/year; however, they are increasing with increasing k and it is quite conceivable that for some high values of k, it will be greater than the average cost. A present worth cost above average cost implies under-estimation of the actual cost (if one is using average cost as a basis for decisions) and *vice versa*. Since the value of present worth cost depends upon the value of k (i.e. carrying costs), it is imperative to have an accurate assessment of the magnitude of these costs. Then only can we get a clear picture of the actual cost. In actual practice very little attention is paid to estimating the carrying costs. Instead, a number is often selected arbitrarily which supposedly accounts for both the opportunity as well as the carrying costs.

To summarize, the present worth approach has some weak and strong features. It does not produce a reorder quantity which is significantly different from the conventional result. Thus, from an inventory operating stand-point there is little impact. But it does give a better picture of the actual costs or profits, as well as identifying those data that may be improved.

REFERENCES

1. Brown RG (1967) *Statistical Forecasting for Inventory Control.* Holt, Rinehart & Winston, New York.
2. Hadley G and Whitin TM (1963) *Analysis of Inventory Systems*, pp. 10–25. Prentice-Hall, Englewood Cliffs, NJ.

RB Misra[1]
AW Wortham
(*August 1976*)

[1] *Department of Quantitative Management Science*
University of Houston
Houston
Texas 77004
USA

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5 No. 1, 1977. Pergamon Press. Printed in Great Britain

International Register of Theses and Dissertations

BOURBONNAIS P (1973) *French: L'utilisation du cycle de la vie comme outil de planification.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

BUTLIN JA (1970) The Effects of Canadian Business Cycles on Canadian Agriculture 1926–1967. MSc Thesis, Ontario Agricultural College. (Available at: the School of Agricultural Economics and Extension Education, Ontarior Agricultural College, University of Guelph N1G 2W1, Canada)

COLE RH (1972) Common Stock Investment in the Franchise Industry. Doctoral Thesis, University of Nebraska. (Available at: the Love Library, University of Nebraska, Lincoln, NE 68508, U.S.A.)

GRUBER S (1971) An Analysis of Agricultural Adjustment Through Multiple Jobholding. MSc Thesis, Ontario Agricultural College. (Available at: the School of Agricultural Economics and Extension Education, Ontario Agricultural College, University of Guelph N1G 2W1, Canada)

HEIN E (1972) *Betriebswirtschaftliche Probleme der Qualitätskontrolle im industriellen Produktionsprozess—Modell einer Qualitätswirtschaft.* PhD Thesis, George-August-Universität Göttingen. (Available at: Niedersachsischen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prizenstrasse 1, Germany)

KOZUMPLIK V (1972) Relative Farming Advantages in Two Regions of Southern Ontario. MSc Thesis, Ontario Agricultural College. (Available at: the School of Agricultural Economics and Extension Education, Ontario Agricultural College, University of Guelph N1G 2W1, Canada)

LE DUFF P (1972) *French: L'organisation et le processus de développement de nouveaux produits dans l'industrie de la biscuiterie.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

MANN D (1972) *Beiträge zur Einbeziehung finanzieller Aspekte in die Theorie der Vermögensbildung und-verteilung.* PhD Thesis, George-August-Universität Göttingen. (Available at: Niedersachsichen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prinzenstrasse 1, Germany)

MULLER-HEINE K (1972) *Analyse agrarpolitischer Ziele und ihrer Einordnung in den gesamtwirtschaftlichen Zielkomplex.* PhD Thesis, George-August-Universität Göttingen. (Available at: Niedersachsischen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prizenstrasse 1, Germany)

NGUYEN T (1974) *French: L'Attitude et le Comportement des passagers aériens.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

SAMSON P (1972) *French: La Détermination de marchés pilotes au Québec.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

SCHAFER V (1972) *Struktur and Probleme eines Simulationsmodells zur Bestimmung eines Werbeträgerplanes.* PhD Thesis, George-August-Universität Göttingen. (Available at: Niedersachsischen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prizenstrasse 1, Germany)

SEBBAN C (1972) *French: Aspects humains des systèmes d'information: Modèle d'implantation.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

WAHNSCHAFFE G (1972) *Die Wirkungen des Monopole auf die Wohlfahrt einer Volkswirtschaft.* PhD Thesis, George-August-Universität Göttingen. (Available at: Nidersachsischen Staats- und Universitäts-Bibliothek, 2400 Göttingen, Prizenstrasse 1, Germany)

WHITE DD (1971) Application of Management Theory and Practice to a Public Mental Retardation Facility. Doctoral Thesis, University of Nebraska. (Available at: the Love Library, University of Nebraska, Lincoln, NE 68508, U.S.A.)

ZWART AC (1973) A Recursive Spatial Analysis of the North American Pork Sector. MSc Thesis, Ontario Agricultural College. (Available at: the School of Agricultural Economics and Extension Education, Ontario Agricultural College, University of Guelph N1G 2W1, Canada)

*The following are obtainable at the Library, Cranfield Institute of Technology, Cranfield, Bedford MK*43 0*AL, UK:*

ANDERSON D (1973) The Influence of Blood Flow Rate on Thrombus Formation as it affects the Design of Mechanical Pumping Systems. MSc Thesis, Cranfield Institute of Technology.

BARRETT RS (1973) An Investigation into the Application of Fluidics to Ship Stabilization Systems. MSc Thesis, Cranfield Institute of Technology.

CHOWDHURY KS (1973) An Investigation into the Problems of Production Management in Multi-National Companies. MSc Thesis, Cranfield Institute of Technology.

DILWORTH SJ (1971) The Analysis of Man–Machine Systems. MSc Thesis, Cranfield Institute of Technology.

DOUGALL M (1971) Production Planning in a Batch Production Environment. MSc Thesis, Cranfield Institute of Technology.

GARNETT RHT (1972) The Assessment of Ore Reserves at Greevor Mine, Cornwall. MBA Thesis, Cranfield Institute of Technology.

GORDON BM (1972) A Search for Replacement Business. MBA Thesis, Cranfield Institute of Technology.

GRITZMAN H (1972) The Market for Computer Bureaux Services and the Marketing Strategy and Organisation of one Bureau. MBA Thesis, Cranfield Institute of Technology.

HEATH JL (1972) A Study of the Out-Patients and Accident Departments of a General Hospital. MBA Thesis, Cranfield Institute of Technology.

HINDLE K (1972) Management and Control of Computing. MBA Thesis, Cranfield Institute of Technology.

HULTON EAS (1972) The Opportunities for Property Development following Siting of the Channel Tunnel Terminus in West London. MBA Thesis, Cranfield Institute of Technology.

HUNT JR (1971) A Housing Decision Model for Milton Keynes Development Corporation. MBA Thesis, Cranfield Institute of Technology.

HUNTER EN (1971) An Investigation into the International Road Transport Industry. MBA Thesis, Cranfield Institute of Technology.

JAMES PW (1971) A Framework for the Preparation of the Strategic Plan in GKN Rolled and Bright Steel Limited. MBA Thesis, Cranfield Institute of Technology.

LAKE J (1971) Fixed Order Cycle Stock Control System. MBA Thesis, Cranfield Institute of Technology.

LATHAM FH (1971) An Investigation into the Marketing of Industrial Products used in the Construction Industry. MBA Thesis, Cranfield Institute of Technology.

LEKKAS AE (1971) Stock Control in a Small Firm. MBA Thesis, Cranfield Institute of Technology.

LINTON TH (1971) An Examination of the Stockboy Inventory Control System. MBA Thesis, Cranfield Institute of Technology.

LOVE JP (1971) Rigidix: A Product Analysis. MBA Thesis, Cranfield Institute of Technology.

MADHUSUDAN M (1970) Further Developments of a Heuristic Method of Job Shop Scheduling. MSc Thesis, Cranfield Institute of Technology.

MEHTA VK (1970) The Application of a Small Computer to the Incremental Control of a Lathe using Stepping Motors. MSc Thesis, Cranfield Institute of Technology.

MILES CJ (1973) The Application of a Visible Record Computer for the Progressing and Control of the Batch/Flow Production of Electronic Sub-Assemblies. MSc Thesis, Cranfield Institute of Technology.

MILES R (1971) Problems of the Interface between Sales Forecasts and Production Programmes at Scott Bader Company Ltd. MBA Thesis, Cranfield Institute of Technology.

MISKIN NI (1971) Lesney Products: A Company in Transition. MBA Thesis, Cranfield Institute of Technology.

MOXON A (1970) The Effective Employment of Professional Mechanical Engineers. MSc Thesis, Cranfield Institute of Technology.

MUIR AJ (1972) A Survey of School Timetabling Algorithms with Reference to their Use in the Computer Compilation of Annual Sports Fixtures. MSc Thesis, Cranfield Institute of Technology.

ORR SW (1970) Setting Hoover Advertising Budgets in Europe. MBA Thesis, Cranfield Institute of Technology.

O'TOOLE BM (1970) The Application of Supplier Quality Control to the Aircraft Industry especially with Relation to Sub-Contracting. MSc Thesis, Cranfield Institute of Technology.

PAGE ML (1973) An Improved Method of Production and Control in a Batch Manufacturing Organisation. MSc Thesis, Cranfield Institute of Technology.

PAPPAS NS (1972) Pilots Manpower Planning in the Hellenic Air Force (HAF). MSc Thesis, Cranfield Institute of Technology.

PARMLEY P (1972) A Mathematical Model of the UK Alcoholic Beverages Market. MS Thesis, Cranfield Institute of Technology.

PEARCE JE (1972) An Application of Group Technology to Batch Manufacture. MSc Thesis, Cranfield Institute of Technology.

PEARCE RC (1973) A Comparison of Vector Loop Methods and Rectangular Co-ordinate Methods in the Kinematic Analysis of Linkages. MSc Thesis, Cranfield Institute of Technology.

PENNOCK DRM (1971) The Problems of Assessing the Market for a Specialist Filling Machine. MBA Thesis, Cranfield Institute of Technology.

PETERSON CI (1972) Production Planning and Control for Small Bore Tube Manufacture. MSc Thesis, Cranfield Institute of Technology.

RAMANI N (1972) A Standard Procedure for the Creation of Computer Files. MSc Thesis, Cranfield Institute of Technology.

RAMSDEN G (1972) Production Programming to Optimise Financial Return. MSc Thesis, Cranfield Institute of Technology.

REIF B (1971) Operational Research Techniques Applied to Urban and Regional Planning: A Survey. MBA Thesis, Cranfield Institute of Technology.

SPIERS MR (1971) Budgetary Control of a Multistage Production Plant. MBA Thesis, Cranfield Institute of Technology.

STANWAY RH (1971) Lawyers in Commerce and Industry. MBA Thesis, Cranfield Institute of Technology.

STEWART JE (1972) AH Allen Ltd—A Market Study. MBA Thesis, Cranfield Institute of Technology.

TAYLOR MH (1972) The Economics of Residential Housing in London. MBA Thesis, Cranfield Institute of Technology.

The following are obtainable at the MIT Microreproduction Laboratory, Massachusets Institute of Technology, 50 Memorial Drive, Cambridge, MA 02139, U.S.A.:

BAXTER CF (1972) A Managerial Examination of the United States Liquid Metal Fast Breeder Reactor Program. SM Thesis, MIT, Cambridge, MA.

BENDER DF (1972) The Approach to Planning Process Formalization. SM Thesis, MIT, Cambridge, MA.

BENNETT HG (1972) Application of Interactive Decision Aids in Agriculture. SM Thesis, MIT, Cambridge, MA.

BOHM RJ (1970) An Empirical Simulation of Estimation Methods for Computing Time-Weighted Rates-of-Return. SM Thesis, MIT, Cambridge, MA.

BOND RO (1970) Flight Frequency Competition on Domestic Air Routes. SM Thesis, MIT, Cambridge, MA.

BOOTH AR (1970) The Use of On-Line Operational Control Data for Management Control. SM Thesis, MIT, Cambridge, MA.

CAMPBELL PA (1970) Managerial Attitudes Toward Corporate Socio-economic Activities. SM Thesis, MIT, Cambridge, MA.

COOPER JB (1971) The Design and Evaluation of a New Distribution Channel for Shipping Services. SM Thesis, MIT, Cambridge, MA.

DEPASQUALE ND (1971) Guidelines for Defense/Space Industry Diversification. SM Thesis, MIT, Cambridge, MA.

DERBY TH (1971) The use of Persuasive Communication Questionnaires to Influence Public Opinion: An Application to the Nuclear Power Plant Issue. SM Thesis, MIT, Cambridge, MA.

FALABELLA G (1972) R & D Task Accomplishment at a US Army Material Command Corporate Laboratory. SM Thesis, MIT, Cambridge, MA.

FAZIO VJ (1972) Consumer Payment Practices and Preferences in the Boston Metropolitan Area. SM Thesis, MIT, Cambridge, MA.

HAAG CC (1972) The Application of Linear Programming Techniques to Airline Cash-Management Decisions. SM Thesis, MIT, Cambridge, MA.

HOUGH LA (1972) An Evaluation of the Major Determinants in Assessing the Capital Requirements for an Educational Loan Program. SM Thesis, MIT, Cambridge, MA.

HOWELL MP (1973) The Supply and Demand for Helium: Case Study of a Critical Natural Resource. SM Thesis, MIT, Cambridge, MA.

HOWLEY JT (1973) Some Implications of Developing National Social Policy. SM Thesis, MIT. Cambridge, MA.

HURWITZ MG (1973) The Replacement of Equipment and the Lease-Purchase Decision. SM Thesis, MIT, Cambridge, MA.

LABRECHE GJ (1972) Pricing and Financing of Transfer Equity in Cooperative Housing. SM Thesis, MIT, Cambridge, MA.

LOPEZ JA (1972) Short-Term Financial Planning in a Multinational Enterprise: A Linear Programming Solution. SM Thesis, MIT, Cambridge, MA.

MARKMAN LE (1972) Management Information Systems: An Application to Family Planning. SM Thesis, MIT, Cambridge, MA.

NAILL RF (1972) Managing the Discovery Life Cycle of a Finite Resource: A Case Study of US Natural Gas. SM Thesis, MIT, Cambridge, MA.

OBERKOTTER HF (1972) Human Resource Accounting in Industry. SM Thesis, MIT, Cambridge, MA.

Vol. 5
1977

RULE E (1972) Consumer Response to Music: Modeling, Measurement, and Management' SM Thesis, MIT, Cambridge, MA.

SAROFIM NS (1972) Leading Indicators: A Tool for Corporate Forecasting. SM Thesis, MIT, Cambridge, MA.

SCHULMAN MG (1972) A Financial Analysis of the Dickey-Lincoln School Dam Project. SM Thesis, MIT, Cambridge, MA.

SHERLING GM (1972) Towards a Formalized Classroom Model. SM Thesis, MIT, Cambridge, MA.

The following are obtainable at the Lyon Playfair Library, Imperial College of Science and Technology, Exhibition Road, London SW7 2BX, UK:

ADAMS SF (1975) A Comparison of Two Recently Proposed Methods for the Travelling Salesman Problem. MSc Thesis, Imperial College, London.

BEASLEY JE (1975) Simula and Simon 75. MSc Thesis, Imperial College, London.

BRISTOW CD (1975) The Unquoted Company in the UK Economy. MSc Thesis, Imperial College, London.

CHAN CE (1975) Determination of a Distribution Channel for an Electronic Door Lock. MSc Thesis, Imperial College, London.

CHIN WP (1975) Quality Control on Market Research Data. MSc Thesis, Imperial College, London.

DIXON AJ (1975) Trends in the Before Tax Distribution of Incomes. MSc Thesis, Imperial College, London.

FORGET PA (1975) Econometric Forecasting Using Survey Data: A Further Study. MSc Thesis, Imperial College, London.

HALL RS (1975) An Application of Graph Theory for Short-Term International Fiscal Management. MSc Thesis, Imperial College, London.

KNOX MP (1975) Programme Budgeting at Teesside. MSc Thesis, Imperial College, London.

LEADBETER PF (1975) Financial Planning Using a Linear Programming Model. MSc Thesis, Imperial College, London.

MASON RLC (1975) Taxation Problems of International Trading Companies. MSc Thesis, Imperial College, London.

NAPAL B (1975) Controlling International Trade in Primary Commodities: Commodity Agreements and Cartels. MSc Thesis, Imperial College, London.

PASTOR L (1975) Stock Performance: Using the Market Model. MSc Thesis, Imperial College, London.

PHUA KT (1975) Progress of the Japanese Computer Industry. MSc Thesis, Imperial College, London.

RICKMAN PK (1975) Bank Branch Location—A Review. MSc Thesis, Imperial College, London.

SABEL JR (1975) The Steiner Problem in Graphs. MSc Thesis, Imperial College, London.

SENIOR DS (1975) Manpower Planning in Data Processing Organisations. MSc Thesis, Imperial College, London.

SPACHIS A (1975) Production Scheduling in a Canning Company. MSc Thesis, Imperial College, London.

STRINGER RJ (1975) An Interactive Corporate Model. MSc Thesis, Imperial College, London.

Information for inclusion in this Section, presented in the appropriate form, should be sent to the Editor of Omega.

BLANK
PAGE

Vol. 5
1977

OMEGA

The International Journal of Management Science
Volume 5, Number 2, 1977

Contents

SAMUEL EILON	107	Editorial—Who Says Production Control is Dead?
	115	Feedback
HUGH FORD	121	Innovation in a Waste Economy
JC HIGGINS, R FINN	133	Planning Models in the UK: A Survey
BERNARD A LIETAER	149	A Planning Model for Underground Mines—An Application in a Developing Country
YORAM ZEIRA, EHUD HARARI	161	Structural Sources of Personnel Problems in Multinational Corporations: Third-Country Nationals
EA LOWE, AM TINKER	173	New Directions for Management Accounting
KR BALACHANDRAN	185	Optimal Design of a Post Office
AL SOYSTER, B LEV, DI TOOF	193	Conservative Linear Programming with Mixed Multiple Objectives
RL FITTS, JB MASON	207	Marketing Segmentation Research—An Application to Bank Services
GL GENTZLER, TM KHALIL, BD SIVAZLIAN	215	Quantitative Models for Optimal Rest Period Scheduling
	221	International Register of Theses and Dissertations

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977. Pergamon Press. Printed in Great Britain

EDITORIAL

Who Says Production Control Is Dead?

"PRODUCTION planning and control has for many years been considered a somewhat glorified clerical occupation, concerned mainly with masses of forms that should be kept in circulation, but having little authority in formulating policies or making decisions: humdrum routine and no glamor. In recent years, however, it has become more and more apparent that production planning and control systems are one of the basic activities that determine the effectiveness of a production enterprise; that problems relating to the utilization of production facilities involve operational analysis and policy evaluations of the highest degree".

Fifteen years have elapsed since I wrote this piece [4, p v], which was perhaps, at the time, a natural reaction against the commonly held view that production and inventory control were fairly mundane tasks, with little to offer in the way of an intellectual challenge. Many text books on the subject reflected this view, and their preoccupation with record keeping and information systems emphasized the need for the production controller first to plan in detail the sequence of operations that jobs have to undergo on various manufacturing facilities and secondly to be able to retrieve the necessary information at any time in order to expedite the jobs through the system and to take action whenever operations deviate seriously from their prescribed schedules. It is against this background that I argued that "there is far more to production planning and control than a few cleverly designed forms", that the decisions involved in planning and scheduling need to be based on analytical considerations of some sophistication, and that such decisions have implications on other functions in the enterprise and on its performance.

In the intervening years production planning and inventory control have flourished, with publications galore on every conceivable aspect of analysis that research workers can contrive. Facets of uncertainty have been explored in great detail (uncertainty in lead times for procurement of goods in inventory control, uncertainty in processing times in production planning, and uncertainty in demand patterns in both), so have problems of sequencing, multi-product scheduling and inventory systems, criteria of performance, production rules, effects of inventory runouts and product variety—to mention but a few of the prominent themes that run through the literature on the subject. So vast, in fact, has the literature grown over the past 15 years, that legitimate questions

have been asked as to whether any further work is really worthwhile, and two powerful criticisms have been advanced in this respect: first, it is argued, the major aspects of production and inventory control modelling have been adequately covered and we are now on a curve of rapidly diminishing returns; not only does further elaboration of existing models promise only very marginal returns to industry and commerce, but the cost of collecting and validating cost and demand data, let alone the cost of implementation of complex control systems, usually outweighs the potential savings or benefits. Secondly, many elaborate models have lost touch with reality and their underlying assumptions are of doubtful validity.

Numerous examples can be cited to illustrate these criticisms. Take, for instance, the EBQ (economic batch quantity) model, one of the oldest in the literature of operational research and which has been rediscovered or resurrected many a time (yet another revival is now afoot, judging by recent papers on the subject), a model which is apposite to certain aspects of the purchasing function but which has been argued to be quite irrelevant for production scheduling; above all, it is a model based on the objective of minimizing unit cost, and it has been shown that the cost function is quite insensitive to the batch-size over a wide range with respect to the optimal solution. And yet, papers continue to appear with various dubious elaborations searching for optimal solutions under questionable assumptions, when the purported improvement, compared with the results of simple and crude models, is of an order of magnitude that does not exceed the level of accuracy with which the cost data can be determined.

Vol. 5
1977

Or take the many elaborations suggested for the production smoothing problem, as if the original work of Holt, Modigliani, Muth and Simon in this field was not elaborate enough. What is the value, the critic may ask, of the intricacies with which the basic model has been embellished over almost two decades? Not that some of the work was lacking in ingenuity or rigour. For example, one ingenious proposition showed that the total cost of the production system can be further reduced if shipments to customers are allowed to deviate from their orders (even when customers are compensated for such deviations)—an interesting proposition that serves only to demonstrate that system performance can generally be improved when the appropriate constraints imposed upon it are removed, but one that has little to do with possible implementation in practice.

Or, yet again, take the famous three-machine problem (scheduling n jobs, each requiring to be processed on three given machines, to minimize total elapsed time), a problem that has defied solution ever since Johnson came up with his neat solution to the two-machine case in 1954. The literature is strewn with various gallant attempts to tackle special cases of the three-machine problem, but how often are we confronted in reality with the need to schedule production through three machines on the assumptions postulated? And if someone does eventually produce a general solution (and inevitably it will

involve elaborate algebra), what reward (apart from the intellectual satisfaction experienced by the investigator) is there in store for industry?

There are other examples in great abundance, but these will suffice. They illustrate the kind of contortions that fascinate the theorists but leave the practitioners cold. As in many other fields, there is a gap between the two camps, and the gap is widening. The dual criticisms referred to earlier, that much of the current work is trivial (because of its marginal potential contribution) or irrelevant (because of its questionable assumptions) is counteracted by the familiar arguments that theory has a value in its own right, that it serves to illuminate new facets of old problems, and that it occasionally leads to practical benefits in unexpected ways. Such arguments are perhaps insufficient as a blanket justification for any kind of theoretical research that may be contemplated, and most research workers will accept that their work needs evaluation, and that consequently some research projects may be regarded as potentially more promising and more beneficial than others, but they do underline the view that many theoreticians hold, namely that the field of production and inventory control is not dead and there is still plenty to explore and brood about.

Another factor that impinges on current and future practice of production control is the advent of the computer, which in 15 years has moved from a peripheral, and in many ways obscure, position to a central role in the manufacturing environment, and this change is no longer confined to large firms, but is increasingly prevalent even in small and medium-sized firms with several hundred employees, where mini-computers in particular have been making considerable inroads. The impact of the computer on the task of data collection and information retrieval has been, by all accounts, far reaching. The computer has taken over the drudgery of keeping and updating inventory record cards, calculation of usage rates and determination of materials requirements, network planning and identification of critical paths for contracts, computation of capacity and resource utilization. This impressive list of activities is particularly significant when many products, components and raw materials are involved, running often into thousands of separate entities that need to be kept under control, or when several complex contracts have conflicting demands on resources.

The impact of the computer on the development of theory should also have been significant: the amount of computation required to find a solution to an intricate model is no longer a bar to exploring it in some detail; there is no need for the analyst to be confined to selecting only the optimal solution with respect to a simple criterion, and an array of solutions to meet multi-objectives can be generated; sensitivity analysis can be undertaken with relative ease, so that the sensitive elements in any given model can be identified and special care exercised in determining their values at the data collection stage; heuristic methods can be more realistically evaluated against bench mark solutions, which are either

optimal or derived from extensive, though partial, enumeration. Above all—and this is perhaps where theory has been lacking—it is possible to compare the merits of different models or approaches designed to solve a given problem, and to assess the relative performance of alternative control tools. For example, it would be enlightening (and of some practical value) to find out whether a particular inventory control system is comparatively more effective than others in certain industries or in managing stores of certain sizes, and it would be equally interesting to determine the circumstances for which the use of linear programming in production scheduling is more promising than alternative methods, and why.

There is no denying, however, that all these theoretical explorations (and many of those listed are overwhelmingly problem-orientated and are clearly concerned with practical issues) are not likely to culminate in earth-shattering discoveries, unless an entirely new approach dawns on the production scene. As intimated earlier, inventory cost systems appear to be fairly insensitive to the control variables that management is thought to have at its disposal, and this insensitivity over a relatively wide range means that room for improvement—at least at the theoretical level—is somewhat marginal. Is it fair to conclude, then, that for all practical purposes—and apart from some enthusiasts who will continue to toy with theoretical models for the sheer intellectual fun of it, cushioned by the comforting feeling that practical benefits, however small, are potentially to be had—production control is dead? That all the ground-work has now been done? That even crude computer systems without too many frills would do more than an adequate job in running manufacturing plants at a high, if not the highest, peak of efficiency?

Vol. 5
1977

On the face of it, the answer must be yes: after two decades of pulsating development, production control has become routine again; the computer has seen to that, and it is in this sense that some may regard production control as dead. It would seem that the computer has done to production control what technology has done to the motor car: it has made the ride smooth and easy, it requires a driver not a mechanic, and the driver needs to follow some simple instructions and not to understand (let alone tinker with) the mechanism.

A mechanism—that is what most computer packages regard the production system as—a mechanism to be manipulated by the storing and handling of information about processing and inventories. But as WA Speck said in his commentary on Swift, "the Universe is not a mechanism, it is an organism", and so is the firm. Have computerized production control systems lived up to expectations in controlling effectively the affairs of such organisms?

It would appear that they have not. A survey recently published by the British Institute of Management (BIM) [5] covers manufacturing operations in 186 plants of varying size, largely in engineering industries; most of the plants produce both for stock and against customers orders, two thirds use computers for various aspects of production and inventory control and 30% even use

computers for detailed shop scheduling. In the light of this apparent level of sophistication, one would expect production to be under tight control, but performance with respect to job waiting time and meeting delivery dates is disappointing: for example in 60% of the plants jobs spent in 1975 less than 30% of their throughput times being processed and more than 70% in waiting, and only 25% of the plants delivered less than half of their customers' orders on time[1] (and this was in spite of the fact that in most cases delivery dates were quoted by the production and inventory control function). Such disturbing performance measures may be somewhat ameliorated by high rates of utilization of plant capacity, but other studies suggest that this is not the case [3]; and in any event, a high level of congestion in the shop would justify a high ratio of job waiting time to processing time, but it is a poor justification for experiencing blatant and continuous violations of promised due-dates (particularly with available modern computing facilities capable of simulating job schedules and of following the expected progress of particular jobs in the system).

It was a quarter of a century ago that Alford and Bangs [1] stated that "the highest efficiency in production is obtained by manufacturing the required quantity of product, of the required quality, at the required time, by the best and cheapest method". Their dictum remains as valid as ever, and one would have thought that the capacity and versatility of modern computers should have enabled managers to take a grip on the situation and to demonstrate that high levels of performance can generally be attained, and not just in a few celebrated model cases. But apparently they have failed, as the quoted performance figures would suggest, and even if evidence were to be presented to indicate that performance has substantially improved over the past two decades (and such evidence, if it exists, is difficult to come by), the fact remains that the present state of affairs is far from satisfactory, and more so in some countries than others.

What is the reason for this failure? We do not have a simple answer to what is clearly a complex problem, which needs to be examined in some detail. There will be those who would postulate that computerization and increasing sophistication of production control systems are quite irrelevant, that the root cause lies in the attitudes and consequent behaviour of operators and managers; there will be those who would question the performance measures, who would argue that vast improvements have indeed taken place, that some slack in the manufacturing plant is inevitable; others would blame it all on the economic system, on the structure of industry and on other exogenous variables.

All these explanations may well be true, but for my own part I would speculate that a major cause lies in our own failure to understand that any production and inventory control system—whether it is manual or computerized—cannot be solely embodied in information storage and retrieval. In 1962 it was the

[1] Similar surveys in the US suggest that manufacturing plants there have fared much better, although direct comparisons between different surveys are not always adequate [2].

paperwork which I felt was becoming the be-all and end-all of production control; fifteen years later, it is still the paperwork, more glamourized, produced more efficiently and elegantly, but it is paperwork nonetheless, since this is what data-processing is all about and this is what most computer packages for production control are designed to do. There are, inevitably, some concessions to modern production and inventory control techniques, such as compilation of status records and computation of certain key parameters as mentioned earlier, but most of these embellishments are natural extensions of conventional data-processing tasks. The paperwork continues now, as it did 15 years ago, to dominate the system, to ossify it. What we have is production and inventory control without control. And no data-processing machinery can be justified unless it is explicitly designed for the purpose of providing management with the means of controlling the system, of taking action to deal with deviations from plans and schedules, of ensuring that desirable performance criteria are met; sheer monitoring is certainly not enough.

Such speculations and sentiments are all the more pertinent when we find in the BIM survey [5] that the overall average cost breakdown for the plant in the sample attributes 51% of the full factory cost to bought materials and components, as against 18% to direct labour and 31% to overheads. Clearly, then, any major effort to control costs should first be directed to purchasing and inventories, whereas industry tends to give priority to the control of direct labour: "Whole work study departments are maintained to control the direct labour content of unit cost and, particularly in the engineering industries, vast amounts of management time and effort are put into negotiations on work rating, job allowances and the like. Yet there are many plants that spend twice as much on purchased material as on direct labour, that do not even attempt to measure purchasing performance realistically, let alone control it" [5, p. 16]. A reduction of 10% in the unit cost of materials will reduce unit cost by 5%, whereas a similar reduction in direct labour costs will have an effect of less than 2%; put it another way: a plant operating at a profit margin of 20% will have its return on capital improved by some 20% (for a given investment, production output and unit price) if the cost of materials is reduced by 10%, whereas a similar reduction in direct labour costs will improve the return by only 7%.

No, production and inventory control are not dead. But they are either taken too much for granted or are being misapplied. A lot of effort seems to be misdirected, and some thought and action are needed to ensure that while our control systems become more expensive, they also become more effective. Both theoreticians and practitioners need to be shaken out of their state of complacency; above all, they need to understand that control without control is abdication to intransigent systems-bureaucracy.

SAMUEL EILON
Chief Editor

Vol. 5
1977

REFERENCES

1. ALFORD LP & BANGS JR (1952) *Production Handbook*. Ronald Press, New York.
2. DAVIS E (1974) APICS state-of-the-art survey (1973). *Prod. Inventory Mgmt.* **4**, 1–11.
3. DUDLEY NA (1975) Industrial Productivity-Scope for Improvement, West Midlands Economic Planning Council, UK.
4. EILON S (1962) *Elements of Production Planning and Control*. Macmillan, New York.
5. NEW CC (1976) *Managing Manufacturing Operations, a Survey of Current Practice in* 186 *Plants*. British Institute of Management.

. 5
77

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977. Pergamon Press. Printed in Great Britain

FEEDBACK

How to Charge for Management Services

BY WAY of preamble, may I explain that I have been involved in the development of useful management accounts for the last 20 years, both as a senior manager and, in my present situation, in a staff role (with, inter alia, management services reporting to me). One of my conclusions is that I am not sure that there is any universal solution applicable to all cases or, for that matter, for all time in any particular case. Charging methods reflect only one aspect of a more complex situation.

The principles which have now emerged in Hill Samuel are broadly as follows:

1. A rapport between users and the management services is fundamental. Successful projects are only likely to emerge when they are developed by an integrated team in which the users have firmly identified their requirements and are satisfied that the proposed technical solutions will meet those requirements in the most effective manner. Inter alia, this means a formal and disciplined approach to every project.
2. Management services must be given defined objectives, and their work throughout must be required to conform to technically professional standards, which should be set down in writing by the head of the department. These will effectively rule out short cuts in the development of projects.
3. A co-ordinated approach between senior company users with the head of management services. In our case, we have established:
   (a) organisational policy statements for each of the service departments, including management services;
   (b) regular review machinery in the form of a Management Services Control Board; its main purpose is to provide a forum for the monitoring of progress towards agreed objectives on a corporate basis, to establish priorities if necessary, review future plans and resource requirements.

The above philosophy makes the actual charging process less critical, because there has been established policy involvement by the main parties affected. It is also probably correct that users apply subjective judgements about value for money, and this factor should probably be taken into account in determining the basis of charging for services.

So far as the spreading of costs is concerned, we apply the following principles:

(a) We separate the computer processing role from project and development work.

(b) We aim fully to recover our processing costs from users, and to establish that they are below comparable rates of a commercial bureau. Recoveries from user departments are based on usage—so if they use the computer less they are charged less. Any under-recovery becomes a central overhead. In fact, management services are then under pressure to reduce costs in line with the under-recovery.
Such a system helps the user departments to be critical of their use of the service. Efficiency within management services, however, is not brought about by such a charging method (nor can it be, I suspect). The usual management

principles must be applied as for a line department, to control costs and improve productivity.

(c) Project work is of a totally different nature. Here we seek a system which encourages sensible use of highly skilled resources but avoids discouragement of their intelligent use by the imposition of unreasonable recharges.

The backcloth is this: we decided this year throughout the bank that we would discontinue charging out all 'overheads' to operational departments, aiming throughout to associate costs with management responsibility. Therefore, we now charge operational departments only with their direct controllable costs and those indirect costs which are subject to significant variation by operational activity.

The remainder of the indirect costs we believe are better identified in total in the management accounts and not charged out on some basis which has to be arbitrary and can at times be a difficult process (since some departments provide services to other service departments as well as operational departments). Apart from the difficulty of arriving at a meaningful allocation, our experience has been that such charges tend either to create more problems than they solve and, inter alia, unnecessary aggravation between operating and service areas or they are ignored by departmental heads on the grounds that 'I can't do anything about that'.

We have, therefore, applied similar principles to management services, notwithstanding that the department does have features which distinguish it from other service departments. To do this, we have identified the more or less fixed overhead element of the project section of management services activity (e.g. accommodation costs, senior departmental management and so on), which we do not attempt to charge out. The other element of cost is essentially variable in nature, principally staff, which is subject to volume adjustment dependent upon user demands. This part, therefore, we do charge out to users. (We also aim to keep a minimum of permanent staff and supplement it where appropriate by making use of external sources but which, when engaged, are controlled by management services in the same manner as our permanent staff). An essential monitoring process is the agreement of a budget for each project prior to its commencement.

KENNETH WINCKLES

Hill Samuel & Co Ltd
100 *Wood Street*
London EC2P 2AJ
UK

IN HIS editorial 'How to Charge for Management Services', the Editor has drawn attention to a very old problem which shows no sign of going away. Whilst it is difficult to quarrel with his rather tentative conclusion about the advisability of treating central management services similarly to other central overheads, it is perhaps worth looking in a little more detail at what should be regarded as 'Management Services'. Most industrial and commercial operations involve both the manufacture (or assembly) of a physical product or service and the marketing of that product and service. Any activity not directly involved in these functions can be thought of as a supporting management service. The management of computing services, personnel services and property maintenance are all recognized as supporting roles and can therefore be treated as central overheads. Management Science workers seem, to me at any rate, to be in a rather different position. Many of their activities can be regarded as 'internal consultancy'; others such as the production and maintenance of models, which are then used regularly as aids to operational decision making, are less obviously 'staff' activities. Indeed, these model building exercises could, in many cases, have been carried out by the line managers themselves. Business Schools and other centres of Management Education are turning out more and more potential managers having the ability to carry out activities hitherto classified as 'Management Science'. It may therefore be worth thinking again about the desirability of having on the one hand 'managers' who know all there is to know about, say, distribution and 'management scientists' who know a lot about distribution and also have the ability to develop analytical or simulation models of distribution problems. Obviously there has to be some cut-off point. Distribution managers cannot be expected to be computer programmers and lorry drivers as well as decision makers. Surely though the ultimate

Vol. 5
1977

goal of these centres of management education specifically concerned with the development of analytical talents must be to see 'Management Science' moved out of the staff side into line management.

J JENKINS

Department of Management Science
Imperial College
Exhibition Road
London SW7 2BX

IT HAS always been my view that when organisations become concerned about whether the charges should be met on central services and regarded as an overhead, or whether they should be paid for by the user directly as and when the services are used, the value of the Group is likely to be marginal. A Group pulling its weight is successful by leave of the reputation which it has throughout the company. If the reputation is high, then treating the costs as an overhead is seen as a worthwhile payment and not as a burden to be carried by the operating divisions. The allocation to the central overhead means one minimises the marginal cost of bringing in management services' advice. We all know that the barriers to achieving acceptance always needs to be kept as low as possible, to overcome the natural suspicion of taking advice on the performance of one's own job from outsiders.

In the past, in organisations in which I have worked, the most satisfactory allocation of these overheads to the operating divisions has been by weighting them on the basis of their proportional use by the division in the preceding year. Actually, the call for work by management services was usually in two categories: category one where to carry them out would involve more than a critical amount of resources and money; and category two which were preliminary studies and small items of work costing little. Category one items would be submitted for approval to a committee on which each of the operating divisions was represented at a senior level. Category two would be decided to be worked on by the head of management services because they would enhance the group's goodwill with the operating divisions; enable them to assess in a preliminary way the feasibility of proposals for work by the operating divisions; and to test the likelihood of support from the operating divisions of work in an area the initiative for which had arisen in the management services group.

Since the job of this committee is to allocate priorities between demands for type one work and to decide the proportion of the total allocated to type two work, it is essential that membership of the committee is senior enough to take the responsibility for making priority judgements. If management services is making a worthwhile contribution to the company as a whole then these senior people will be aware of their value. They will not be worrying about the allocation of the costs of the services in too strict a way.

RH COLLCUTT

Manchester Business School
Booth Street West
Manchester M15 6PB
UK

Planning for Freedom

ALTHOUGH in Planned Economies Governments have hoped for certain industrial sectors to proceed 'immune from the vagaries of the market place' [2] they have demonstrably failed. Many of the 5 year plans developed by countries for the economic growth seemed to have been based on a closed system analysis ignoring the effect of world markets. Similarly a business enterprise must be aware that discontinuities may occur in the environment (e.g. oil crises, rapid change in currency rates) that can make the corporate strategy irrelevant.

On the other hand, to be responsive only to the whims of the consumer must lead to disaster. Planning must attempt to influence the market and maintain the organisation in a posture that allows it to respond to change in the markets it cannot control, i.e. offensive and defensive planning. The degree of planning required will depend upon the volatility of the market and the size, structure

and location of the enterprise. The greater the volatility of the market the greater the emphasis on the size and structure that allows rapid adjustment [3].

Planning within an enterprise has consequences for the freedom of action of its members. The constraints and opportunities that the implementation of a plan involves is considerably influenced by the management philosophy (political–social system) that is characteristic of the enterprise. The structures that evolve from the management philosophies can be likened to the planning systems described by Van Gigch [4]; they can also be described using the same attributes such as rationality, equality, experts and elites, participation and control.

For the purposes of this comment a simple model may be useful in explaining some of the consequences for freedom of planning within an organisation.

Planning which involves the identification of objectives and the setting of targets may seek to elicit a wide range of contribution from members of the enterprise or be restricted to experts and elites. The target set may thereby be perceived as imposed or agreed.

The attainment of the targets requires the exercise of some control which is translated into operating procedures which inevitably reduce the choice of behaviour available to members of the enterprise. If such procedures are unduly restrictive their observance may be seen as constraints to individual freedom thereby inducing frustration and aggression.

Research which has been undertaken on emotional arousal associated with frustration suggests that if it is prolonged there is a reduction in the level of performance of the individual concerned [1]. Attention is diverted from the tasks to be completed to the imposed constraints. The lowering of performance can lead to an effort after greater control by top management which imposes a further reduction of choice of action through the operating procedures. A vicious circle is thereby set up. The system has become punitive.

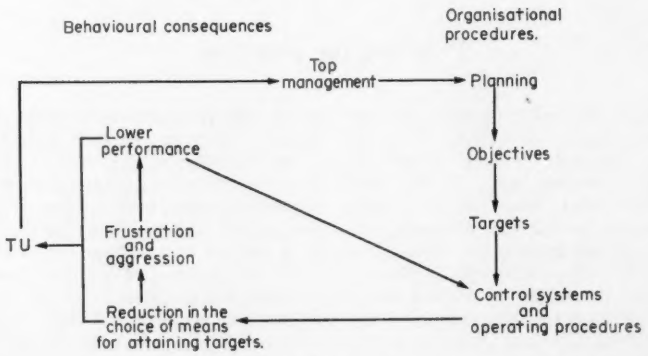
If management do not appreciate the processes causing the lowering of performance and modify control systems accordingly members may seek to go outside the enterprise to a trade union. Direct access to the top of the enterprise by the trade union can result in modification of the operating procedures which generated conflict. Whether or not this would lead to an improvement in performance depends on how entrenched the positions of the conflicting groups have become.

From the above considerations it may be concluded that whilst planning is an essential part of any management practice it must take into account the obstacles and opportunities that affect members of the enterprise in attaining targets that have been determined. The approach to planning adopted may bring with it obstacles which generate conflict and reduce performance or opportunities which encourage achievement of organisational objectives.

Vol. 5
1977

REFERENCES

1. COFER CN and APPLY MH (1967) *Motivation; Theory & Research*, Chapt. 8 & 9. John Wiley, New York.



2. EILON S (1976) Planning and freedom. *Omega* **4**, 501–505.
3. LAWRENCE PR and LORSCH JW (1967) Organisation and environment: managing differentiation and integration. Harvard Business School, Boston.
4. VAN GIGCH JP (1976) Planning for freedom. *Mgmt. Sci.* **22**, 949–961.

C BAKER

Durham University
Business School
Palmer's Garth Hallgarth Street
Durham DH1 3LB
UK

1. 5
77

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977. Pergamon Press. Printed in Great Britain

Innovation in a Waste Economy[1]

SIR HUGH FORD

Imperial College, London

(Received October 1976; in revised form October 1976)

Despite the creativity of British engineers, British industry continues to lack innovative drive because of a climate inimical to commercial success. Contemporary industrial economies, unlike that of Victorian Britain, are geared to produce rapidly consumed and discarded products. Victorian engineers sought primarily to add to society's stock of capital rather than to consumption, and their designs were intended to endure rather than to promote change. All were subject to scrutiny on strict commercial criteria as a measurement of their fitness for their intended purpose. This represents a change in what society demands of the engineer. The innovator is, in Britain, subject to particular discouragements in the form of: low prestige attached to engineering, inadequate financial provision for investment, high corporate and personal taxation and bureaucratic encumbrances by government. By contrast, to flourish, innovation requires leadership, reward, recognition, a flexible economy and a flexible workforce, and adequate and cheap capital.

l. 5
77

ON EVERY aspect of the innovative process, an alarming amount has been written, discussed at Conferences and Seminars, case histories analysed and questionnaires distributed.

So it is unlikely that much can be added that is wholly new or that will help the world of engineering to be more innovative or to help the innovator to help himself. It is all there, in the published work—some of the best in the proceedings of the British Association: all I can hope to do is to summarise it and, from the viewpoint of one who has tried hard to innovate for 36 years or more, to answer this question:—

"Why, since so many people have written about innovation and analysed the successful and unsuccessful attempts at innovation, have indicated its nature, its needs, its consequences, why, I repeat, is the United Kingdom apparently unsuccessful in innovation?" Knowing what are the ingredients for success, it is remarkable that so much erudition has not produced results or turned those who have made the studies into millionaire innovators acclaimed by the world for the benefits they have bestowed upon Humanity. Perhaps the explanation lies in the reflexion that the race-course tout and the investment advisor do not

[1] Based on the author's Presidential Address to the Engineering Section of the 1976 Meeting of the British Association for the Advancement of Science.

make fortunes out of their knowledge of the form they proclaim with such certitude.

The interesting fact emerges, however, that the papers I have found to be the most valuable and arresting have been written by innovators, though they would be quick to tell you they are not rich men. If, as I believe, the ingredients for success in innovation are well known, the cause of our poor showing may lie elsewhere than in the creative minds that these people possess. I hope to show that innovation can only be realised in an engineering context in the right environment. It is my observation that Britain's engineers are today more creative, more productive of new, exploitable ideas than they have been for a long time, and I am led to the conclusion that it is not a lack of ideas, enthusiasm or ability to realise innovations but that the climate is such as to prevent the innovation even though it has all the ingredients for commercial success.

SOME DEFINITIONS

By 'innovation' I mean the carrying through of something new (a new idea, a new application of an existing idea, a new product or a service) to the point of commercial success. In this context, an invention is a creative step, potentially innovative: there are many inventions but few stand the test of fitness for purpose in the market place. I am also confining my remarks to innovation in engineering, and since my life has been spent predominantly in mechanical engineering it will inevitably colour my remarks. The engineer is not to be categorised with the scientist: the scientist concerns himself with the pursuit of natural knowledge and the formulation of its laws by the exercise of the scientific method. The engineer uses scientific knowledge and the corresponding technologies to useful ends, in the contriving of machines and works fit for their purpose and acceptable in the market place. The criteria for good science and good engineering are different, therefore, for it is not only necessary to develop new ideas, concepts or facts but to apply them and exploit them in a real situation. To say that "the operation was successful but the patient died" can, in the context of medical science have a dash of meaning to it. For an engineer to say the invention was successful carries with it the implication that it not only was a bright idea which in physical terms realised its creator's vision but was acceptable as fit for its purpose in the market place and gave an acceptable return on the investment of time, people, money and material resources.

A WASTE ECONOMY

I have asserted that innovation has to be judged in terms of the environment in which it is both carried through and utilised. By the title "Innovation in a

Vol. 5
1977

waste economy" I am taking the view that we live in a society in which we are not concerned with the creation of wealth so much as the creation of waste. By a waste economy I mean an industrial society in which the economic forces are predominantly influenced and controlled by large production of rapidly consumed commodities. Inherent in this view point is the observation of the short life of many things that are produced; of design for obsolescence; or for a fixed useful life at which point you 'trade it in' or 'throw it away'. The justification for such a system is a very wide distribution of goods at low initial price. This is the environment in which the engineer today finds himself and from this state of affairs many other consequences and results flow.

At this point I wish to establish that the ability to innovate and the style and kind of innovation that is possible depends upon the boundary conditions under which the innovator has to work.

The frame-work in which the engineer has been required to innovate has changed with time. The great period of creative engineering and innovation undoubtedly took place here in Britain towards the end of the 18th century and during the 19th century. Yet it is true to say that the Victorians believed that their form of society, and particularly their possessions, were secure against change. The paradox is that this same outlook was reflected in their engineering; for the things they conceived and invented were seen, not as the vehicles of change, but as fit for their purpose and of service to the society into which they were born. They saw the development of motive power and of machines as a means of increasing total wealth and to allow society to enjoy life more abundantly within the framework as they found it. Not only were the products intended to last but also the designs from which they were made were not expected to change. The existence of a large and skilled work force provided the means of maintenance and repair and for steady improvement. The great innovative steps that were made by the early engineers were not intended to lead to an industrial revolution but to an evolutionary process which added to wealth.

This attitude is well illustrated by the great George Stephenson himself. Samuel Smiles, in his biography of the Stephensons, repeatedly shows that George was both concerned that the enterprise and risk-taking by his clients and the benefit to the public at large should not be jeopardised by the ignorance of Parliament or the actions of vested interests. At one point Smiles says "To this it must be added that in urging these views, Mr. Stephenson was strongly influenced by commercial considerations. He had no desire to build up his reputation at the expense of railway shareholders, nor to obtain engineering éclat by making 'ducks and drakes' of their money. He was persuaded that, in order to secure the practical success of railways, they must be so laid out as not only to prove of decided public utility, but also to be worked economically and to the advantage of their proprietors—in fact, commercial speculations. He therefore endeavoured to render them financially profitable; and he repeatedly

declared that if he did not believe they could be made to pay, he would have nothing to do with them" [9].

In this quotation, I submit, there is every aspect of good innovative practice in an engineering context; the engineer innovator in command and concerned not only to realise his ideas but to do so in the public interest, profitably and economically, and in a free economy, yet within the boundary conditions of the then situation.

Today the cry is for change, an attitude that reflects the social and economic revolutions of the present century. It really started, in my view, towards the end of the last century because of the lack of a large skilled labour force in America which demanded that more components and commodities had to be made with a small skilled work force. This, coupled with the great population explosion in America and the need for transport over large distances, demanded a prodigious internal productive capacity that could not be met from overseas. I believe that Britain has absorbed much of the thinking and action of a consumer society, but has retained an instinctive preference and prejudice in favour of a much longer time scale. In mechanical engineering our exports are still predominately in the range of products we have traditionally made well, the heavier capital goods market and the more durable of the consumer durables. Our manufacture depends upon a high level of employment and until the last 20 years, used a high proportion of skilled labour. I suspect that in the longer term our national instincts are right. A waste economy depends upon cheap raw materials in abundance, and the availability of continually increasing power per worker. Couple this with the belief that because scientific knowledge allows some new technological advance to be made, then situations must be changed to make use of it, and the end point is inevitable: advertising and enthusiasm does the rest.

Vol. 5
1977

The large scale development of plastics is a good example of this. There are real uses for the bulk polymers, but these are for relatively small quantities. The fact that by 1980 the total volume production of plastics is predicted to exceed the volume of steel; the reflection that they are predominantly used for packaging or for containers or components of very short life, is proof of the total change in what the world public *thinks* it wants—whether you judge this to be real or artificially created because the plastics are there and can be made profitably and in large quantities. Consumer products, and particularly those used in the home, are now frequently scrapped when a component fails and this apparent extravagance is made possible by new production engineering methods that permit the cheapest assembly of readily available materials.

This overwhelming demand for change in large 'throw-away' markets to my mind leads to a contradiction. If we consider for a moment the manufacture of automobiles, the life now looked for is about 5 years with the sales force expecting to have virtually a new model each year. Yet frequent obsolescence of a major kind, which should be the consequences of this continual pressure for

change coupled with 'trade in' and 'throw-away' philosophy, is the exact opposite of what is needed in large scale mass production manufacture. The motor car industry prefers and requires an evolutionary development because it has to avoid large capital expenditure and must build reliability and safety into its vehicles. This is not only true of the motor companies themselves but of the great number of component suppliers whose task becomes impossible if there is a continual demand for changes that may well mean very expensive modifications or even scrapping of production lines of engine parts, brakes, gears and so forth. For the car body companies, even a change of styling can mean vast expenditure on new press tools.

Despite all the advertising and marketing noises, all the emphasis on the new model, it is remarkable how slowly the essential features of a car can change.

In the annual Brighton run for veteran cars there is a four cylinder Mercedes sports car made in 1902 that was built to do 70 miles an hour and is still more or less capable of doing this. Now this is the fastest speed that one can legally travel on the roads of Britain and I am led to wonder whether it makes sense to demand change for its own sake—and the terrifying waste of resources and raw materials—merely to make some small change which we are led to believe the market requires. In terms of fitness for purpose—to get from *A* to *B* in suitable comfort, reliably and safely and with the minimum overall cost during its useful life—I suspect a motor car would be different from its present form and specification.

I have emphasised this aspect of the waste economy because the engineering designer and the innovator has to work within the framework in which he finds himself. At least, this is the belief in 'the customer is always right' and ideally, in this case, the minimum economic life of a product has to be predicted from the cash flow calculations applied to market assessments. As Parker [8] has pointed out "unfortunately these analyses are not yet reliable. For example, Alan and November [1] examined a number of project case histories from seven firms and found that markets were, on average, overcast by a factor of three". [8]

The notion of a throw-away society depends upon the unlimited supply of raw materials of all kinds. It further depends upon the desire of people to work to produce more and there are few politicians who do not subscribe to the view that it is necessary to have an expanding economy for survival. Yet it seems to me that there may be some contrary forces already coming to play that may change the boundary conditions under which the innovator has to perform. The almost religious belief in the virtue of hard work of Victorian times is rapidly disappearing. The increasing lobby for safety and reliability, the avoidance of effluents and the destruction of the environment, added to the greater cost and scarcity of raw materials may well require some drastic new thinking. However, so long as we retain the present framework in which we have to innovate, those of overwhelming importance are related to economic factors of one sort or another.

In addition to the limitations placed by a waste economy, the other important factors are:—

1. Government action and policy, particularly its taxation policy.
2. The level of reward for risk-taking in the introduction of a new product.
3. The form taken by statutory requirements for the presentation of company accounts and performance.
4. The Civil Service.
5. The high cost of money.
6. Competition, particularly in relation to overseas markets.
7. The time scale.

All these factors add up to saying that successful innovation requires a suitable social, economic and industrial environment.

As Sir Alec Cairncross has put it: "The innovating manager in Britain has a hard time of it compared with other countries. He has to contend with government regulations that absorb his time, with restrictions imposed by his workers or by his competitors, with public apathy, with the slowness with which things get done. Innovation is a form of investment and when the economic and social climate is uncongenial to investment it is naturally discouraging to innovation too. If the whole community was dedicated to economic growth and all that it involves I should expect the rate of innovation to respond automatically and to sustain faster economic growth. But I cannot help suspecting that interest in growth is fainter than we pretend" [2]. Sir Alec also emphasised the apparent lack of correspondence between expenditure on scientific research and the ability of the Country to innovate and gain economic advantage from its endeavours.

He went on to say "Governments tended to draw the conclusion that heavier expenditure on science would yield a dividend in GNP and their emphasis on science increased as time went on. But the basic weakness lay in engineering in the lack of appreciation in British industry of the value of outstanding design engineers and the simultaneous lack of interest in the schools in an engineering career".

It should also be borne in mind that buying a licence or knowhow is a trading expense and involves no risk and that this latter point is particularly important in Britain, because heavy taxation has engendered a distaste and suspicion of new projects and ideas. There is an extremely large 'hump' of resistance to the

acceptance of something new until it can be seen in operation elsewhere, especially if the elsewhere is Germany or America.

In a remarkable paper to the Institution of Mechanical Engineers, Jones has pointed out the great importance of the way in which we assess the health of a company as to both its ability to finance itself and to create new ideas and products [4]. He has contrasted the difference between the British requirement for an Annual Balance Sheet and Profit and Loss Account with the Japanese system in which the Company has to show, twice yearly, not only financial information but operating information as well. The Japanese have the fastest growth of output per employee in manufacturing industry in the world and they have also developed measures of industrial efficiency upon which judgements can be made about the effectiveness of their industry, that is, the added value per employee and per unit of capital employed. The insistence in the British system on return on investment has produced an artificial state of affairs and Jones shows that although we appear to come out well compared with the Japanese on the wealth created per unit of total assets employed it is only done because of the very low allowances for depreciation and re-investment in better plant and methods. I would like to refer you to his paper for the quite remarkable facts and statistics which he quotes. Averaged over several hundred manufacturing companies in Japan and Britain, he shows that the total assets per employee in Japan is around £28,000 as against £9000 for his equivalent in Britain, Moreover, the added value per employee in Japan is about £7400 as against a British figure of £3038.

Jones then takes the case of large nationalised and semi-nationalised industries and compares (1) the British Steel Corporation with New Nippon Steel Company and (2) British Leyland Motor Corporation with Toyota. The statistics relate to the performance of the Companies within their respective countries, that is, stripped of overseas divisions and associated companies.

For the steel industry, the remarkable facts are that British Steel Corporation employs 2·8 times as many people as its Japanese counterpart with about £7000 of total capital per employee as against nearly £44,000 available to each employee in New Nippon Steel. Again, sales per employee are about 4·2 times as great in Japan as they are in Britain while the added value per employee is nearly three times as great. Of the total added value, (that is, of the wealth created) the British employee takes 75·1% where his Japanese counterpart takes 47·3%. The consequence shows up dramatically in depreciation where British Steel Corporation, despite its favoured position in borrowing money, retains 11·1% of total added value where the Japanese Steel Company is putting away 21·5%.

The position is even worse with BLMC. It employs 3·92 times as many people as Toyota although its sales per employee are only 18% of the performance of Toyota. The added value at £2829 per worker is only 37% of that in Toyota but the most telling figure is that the employees in BLMC take out 91·6% of the

added value in wages and social services, leaving virtually nothing for re-investment, and for all the other needs of a dynamic business, or even for the tax collector. It should also be noted that the depreciation per employee in Toyota is more than eleven times that in BLMC.

I have quoted these figures at some length because they highlight better than anything else the terrible burdens and handicaps that the innovator faces in the UK. Innovation requires investment, financial incentives, development of new production lines and methods, a work force willing and able to accept changes in employment patterns, a market that can respond to the innovation against a sound financial judgement of its potential.

Sound financial judgement about new ideas is unlikely where industry is being run virtually for the benefit of the workforce. BLMC is not alone in showing figures of this kind, indeed I know of cases where the employees were taking more than 100% of the added value. Innovation is impossible in such a climate.

The situation is further exacerbated by Government attitudes or lack of appreciation of the real situation. Where Government policy is to borrow enormous sums to finance current expenditure we are eating up our substance that should go to reinvestment in better plant, better ideas and products. Such a policy can only make worse our position as against our competitors. There is nothing available to invest in new ideas. It is not surprising that the innovator finds inevitably that he has to go abroad to find the support he needs to realise his innovation. Not only that—as I have found to my personal cost—the innovator has to dissipate his time and resources (as does the nation as a whole) in combatting the increasing number of public employees, to deal with the growing spate of pointless requests by officials and bureaucrats for data and statistics on every conceivable thing. Indeed, the smaller the company, the bigger the burden this becomes at a time when all the effort is needed to bring the enterprise to success. In one project in which I am concerned the company has nearly 50% of its staff involved in recording information af all kinds and relaying it to a whole army of officials. I said at the outset that on my observation we are more imaginative in a commercially realistic sense than we have been for a very long time and yet the frustrations and the barriers to the wealth creating activity are greater than they have ever been. I find this situation extremely disturbing and Governments of both colours are equally to blame for the steadily increasing load of bureaucracy and restrictions that the wealth creators have to shoulder. No political party seems to be really aware of the inevitable end point of their policies.

These burdens might not be too bad if the innovator, in return for his initiative and unrelenting hard work, saw a reward at the end of it but he can see little joy in this regard. In the last 4 years, as Lord Nelson has pointed out [7], the Senior Manager has seen his standard of living fall by 33% and his position relative to equivalent people in central and Local Government has deteriorated discouragingly.

Vol. 5
1977

What does the innovator need?

Innovation (and I should repeat that by this I mean the successful introduction of a new product or service) requires a creative genius, disciplined to the needs of the market place. It is a rare quality and needs fostering, encouraging and rewarding if it is to create the wealth we need. The innovator needs not only the right environment in which to work but innovation depends enormously on the people who innovate. It has often been said that James Watt had the very good fortune of meeting Matthew Bolton who was able to see the great benefits and advantages to be gained from Watt's inventions and to cope with all the problems of innovation. This has been very happily summed up in the following words:— "It seems that the very factors that make an inventive mind—an impatience with things as they are, a tendency to look at them from an unusual point of view and a spirit to challenge accepted authority—are likely to make that mind unsympathetic to the methods that have to be employed by administrators when the invention passes to the stage of innovation when numbers of other men have necessarily to be brought in".[5] The importance of the organisation for innovation cannot be over emphasised. This aspect of the innovative process has been adequately dealt with in several papers. A very good account occurs in [3]. Again, Langrish *et al.* [6] and others analyse the relative importance of various factors in the innovative success of firms in various kinds of industry and show that the most important single factor in success is the presence of an outstanding man in a position of authority, either in providing the basic idea or by backing and promoting the idea produced by somebody else. For most of engineering industry, the 'important individual' is responsible for over 40% of the successful innovation; on the other hand the clear identification of a need through market studies and the like is relatively unimportant. While it may be fair to say that a minor innovation is often the direct result of looking at a market need, major changes predominately have come from an accidental discovery followed by an innovator recognising the potential usefulness of the accident.

All this adds up to the conclusion that while innovation is dependent upon a market and upon the development of something fit for its purpose, the innovator needs above all an enthusiasm and a dedication to purpose that overcomes all the opposition and set backs, the ability to handle people and resources and to make the cash available and control it over the whole span of the introduction of the innovation. None of this should be construed as meaning that the nature of the innovation is unimportant. Of course it is overwhelmingly important to have answered the question "What is really worth doing?" and to have used all the disciplined analyses and criteria for successful realisation of a new idea. The innovator, of course, needs first and foremost the new idea, but the best idea in the world will flounder unless it is in the charge of a man of courage, tenacity and enthusiasm who is determined to see it through.

Why do we innovate?

In the first place I believe that there are some people who cannot help themselves and I suppose one would call this the need for personal satisfaction in the realisation of an idea; but I do not think this is enough. I believe that the innovator needs to satisfy not only personal ambition and justification but should expect to be able to look for a reward for providing something which is better than that which already exists. Whether he works as an individual or as an employee of a company, there should be both a corporate and a personal profit for the intense effort the innovator puts into his work. Many may be motivated by an altruistic desire to serve mankind but "To each according to his needs, from each according to his ability", may sound fine but has never brought home the bacon. Russia is not outstanding for its innovations.

It is my personal belief that the innovator requires a free economy for real success. By this I mean that the society in which he operates must be capable of responding to the introduction of new ideas where both the innovator and those who will see a benefit from the innovation can freely cooperate and expect to look for some profitable reward. A monopolistic and monolithic system has too many vested interests, too many commitments to its large scale objectives to be able to accept changes easily or economically. Where the ponderous workings of bureaucracy have finally committed resources on a certain course, the time scale is likely to be a long one and the innovator will be bankrupt or dead before he can hope to influence so rigid a system. Time is of the essence in innovation. World competition in activity and ideas is such that the innovator can be sure that there are several people close on his heels or even ahead of him and the hump of resistance to the acceptance of his product can be the most destructive element of all.

The innovator also needs cheap money. At the present time this is something he very rarely gets though here the presence of an outstanding man in a position of authority (which must inevitably mean the control of the cash) can give the employee of a large company the equivalent of cheap money. However, the money still remains expensive to his company and in the long term the innovation must be such that it repays the investment. However, it does mean that the servicing of the money and the risk involved can be delayed. It has to be recognised that a great part of the innovation here and in the United States is the result of the relatively small enterprises and to these the present high cost of money is very serious.

The way ahead

It would seem from what I have said that everything is steadily building up to oppose successful innovation and that since we are caught up in a Socialistic Welfare State, the end point is inevitable and that in the face of a monopolistic

Vol. 5
1977

and bureaucratic system, the innovator can fold up his tent and silently steal away. Having been concerned with innovation in mechanical engineering over the past 36 years, I find myself frequently subscribing to this view. But, somehow the creative spirit, the enquiring mind, coupled with the practical ability to see imagination channelled into a successful endeavour will always motivate a small handful of people even though the illusion of gaining a reward for one's endeavours may disappear in the hard face of reality. There is no doubt that Britain's economic position could be enormously helped by an attitude of mind towards the value of innovation and a framework in which innovation was properly rewarded. The important thing is to try to recognise what the ingredients of the framework should be to bring about this result.

Undoubtedly, the problems of wealth creation in general have much in common with the problems of the innovator.

A healthy atmosphere in which industry can flourish will automatically help the flowering of the innovative mind. But I think something more than that is required. I believe the innovator requires special recognition and reward for his efforts. I believe that there is a class of activity where special rewards could be recognised and given and I have proposed this many times in many places over the past twenty years. Innovation covers a very wide range of activities and contributions. Much of it is the result of a lot of people interacting to produce a better result; on the other hand, invention, know how, experience, knowledge, and advice springing from experience and knowledge, can be identified whether in a company or a team of people or an individual. It can, I submit, be traced back to its source. For example, British consulting engineers contribute around £250,000,000 per year to the balance of payments in return for advice, designs and projects. In terms of added value per employee and per ton of raw materials imported into Britain this is a very handsome return quite apart from the considerable amount of exports of hardware of all kinds it engenders. Again, royalty payments and licences on inventions and developments or consultancies arising from specialised knowledge are traceable to those who have generated them. My proposition is that all such overseas licence fees, royalties, invention payments, consultancies should be free of tax to the organisation or the individual giving rise to them, providing that they are brought back to this Country.

I believe that this would help significantly to tip the scale in favour of the innovator and make very little loss to the UK Treasury. I believe it is necessary, also, to allow the costs of innovation to be written off as an operating expense not only as regards the initial stages of research and development but also the much more expensive process of marketing the project, of overcoming customer resistance and of tooling up for production. Far too many good projects have failed because of the very high costs of introducing them into the market and providing the inevitably higher after-sales service of a new product.

There has been increasing evidence lately of the resistance of the work force to the introduction of a new product because it requires less people or different

skills. The high employment level for a given output in Britain compared with our major competitors is one consequence of this kind of thinking. One device that is used when the work force refuses to allow an innovation is to start up in a new factory with new employees elsewhere. This is a wasteful device and may well kill an otherwise successful innovation. It takes time and time is not available to the innovator.

I believe that there is a need for cheap risk capital for innovation.

Government can borrow £10,000 m per year for recurrent expenditure with very little to show for it at the end. Yet it cannot make available a fraction of that at low interest rates to provide the life-blood of wealth creation in the longer term or forego its taxes that wither the healthy growth at the moment when it needs sympathetic nurturing.

The burden of my song is then that there is nothing so much amiss with the British innovator that the right environment could not put right. He is as good as he ever was: given a better financial climate, greater freedom to exploit his ideas and a better reward for his enterprise and hard work, I believe he can rescue us from our predicament: without these we are in a sorry plight indeed.

REFERENCES

1. Allen DH & November PJ (1969) A practical study of the accuracy of forecasts in novel projects. *Chem. Engng* **229,** 252–262.
2. Cairncross Sir Alec (1971) Government and Innovation. Presidential Address, *Meet. Br. Ass. Adv. Sci.*
3. Jewkes J, Sawers D & Stillerman R (1969) *The Sources of Invention,* 2nd Ed. MacMillan, London.
4. Jones FE (1976) The economic ingredients of industrial success. *Proc. Inst. mech. Engrs* **190,** 115–120.
5. Jones RV (1974) Engineering creativity. *Proc. Inst. mech. Engrs* **188,** 585–596.
6. Langrish J, Gibbons M, Evans WG & Jevons FR (1972) *Wealth from Knowledge.* MacMillan, London.
7. Nelson Lord (1976) Chairman's Statement to Shareholders AGM of General Electric Co.
8. Parker RC (1970–71) The art and science of selecting and solving research and development problems. *Proc. Inst. mech. Engrs* **185,** 879–894.
9. Smiles S (1874) *The Lives of George and Robert Stephenson,* Folio Society Edition, p. 233. London (1975).

ADDRESS FOR CORRESPONDENCE: *Professor Sir Hugh Ford, Imperial College of Science and Technology, Department of Mechanical Engineering, Exhibition Road, London SW7 2BX.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977. Pergamon Press. Printed in Great Britain

Planning Models in the UK: A Survey

JC HIGGINS

R FINN

University of Bradford

(*Received October 1976*)

This paper presents the results of a survey on the use of computer-based models in the broad areas of planning and strategy formulation in UK companies. This survey updates and complements other recent surveys, which exclusively considered the use of corporate models, and shows a continued increase in the numbers of these models in use. In particular, models in the following broad categories were examined: forecasting, marketing, personnel, production and models of 'one-off' situations, as well as corporate and financial models. For such models the paper includes details of: frequencies of occurrence and relationship to company size; professional roles of the model-builders and users respectively; frequencies of use. In these and more subjective ways, the survey attempts to assess the place and value of computers in planning.

1. 5
77

INTRODUCTION

IN THE last few years, several surveys have been conducted [1, 2] which have investigated the use of corporate models in business organizations. However, the other types of model used in planning and decision-making areas have been largely neglected by the surveyors. It is the objective of this paper to present the findings of a recent survey which covers the broad area of the use of computer-based models in the planning function of business.

This survey, because of its broader nature, does not address the topic of corporate models in the same depth as the survey by Grinyer and Wooller [1] but attempts to monitor developments since that survey. Furthermore, because a random sampling method was employed, we have been able to examine the views of those companies who at the moment do not possess a corporate model as well as those of the companies who do.

The results presented here are part of a wider research covering a random sample from the UK companies listed in "The Times Thousand Companies". Two postal questionnaires were addressed to a number of companies and some 56 took part. The industrial background of the respondent companies is

summarised in Table 1. The first of the two questionnaires was intended for the chief executive and examined senior management's attitudes to quantitative techniques. The second questionnaire was to be completed by the executive responsible for planning and had a dual purpose. In part, this was to gather background information to supplement the first questionnaire but its main objective was to examine the structure and practice of corporate planning in the UK. In this paper, the emphasis is on the use of computer-based models in planning. In addition to the postal questionnaires, a number of interviews were conducted in order to give greater depth to the research.

TABLE 1. CLASSIFICATION BY INDUSTRY OF COMPANIES PARTICIPATING IN SURVEY

Industry	Number of companies
Agricultural supplies	1
Banking and finance	3
Beers, spirits and tobacco	2
Building and construction	3
Ceramics	1
Chemicals	2
Cinemas, theatres and television	2
Clothing and footwear	1
Coal, nuclear energy and steel	1
Communications	1
Data processing and office equipment	3
Electrical and radio	2
Engineering and metal	5
Food, catering etc.	6
Glass products	2
Household products	2
Industrial plant manufacture	1
Insurance	2
Motor	1
Newspapers and publishing	2
Packaging	3
Services	3
Shipping	1
Textiles	1
Transport	3
Other[1]	2
Total	56

[1] Two companies did not identify themselves or the nature of their industries.

Vol. 5
1977

COMPUTER-BASED MODELS EMPLOYED

Computers are employed extensively at the operations level of most organizations today, either as data processing devices or as the tools of such specialists as operational researchers. The use of computers and mathematical models has developed more slowly at higher levels of management, partly because problems at this level are not so well-structured. However, with the spread of corporate planning the demand for more analytical tools has increased and the computer has been employed in such areas as forecasting, marketing and corporate modelling.

Most large and medium sized companies will have some sort of computer facility, but not necessarily available to all their departments. Table 2 gives a summary of the computer facilities available to the planning departments of the companies taking part in the survey broken down by company turnover.

Batch processing still constitutes the main method of access to the computer with 61% of the planners using this form of job entry. It is noticeable, however, that as the company turnovers increase, so does the importance of conversational access relative to batch processing. In the lower two turnover groups, batch processing is more common than conversational mode; in the third group, both types of access are equally common; in the highest turnover group, conversational systems are more common than batch processing systems. Two other points are noticeable in Table 2:

(i) Not unexpectedly, the range of computer facilities available appears to increase with company turnover,

TABLE 2. COMPUTER FACILITIES AVAILABLE FOR PLANNING

Facilities available \ Turnover groups £M	15–20 (%)	20–100 (%)	100–300 (%)	>300 (%)	Other companies[1] (%)	Totals (%)	
Batch processing	7	22	5	15	12	61	means of 1·51 systems per company
Conversational	5	5	5	22	10	47	
Remote batch	0	7	3	10	3	33	
Conversational Remote batch	0	0	3	7	0	10	
None	5	7	5	0	0	17	
% of total represented by each group	15	29	15	27	14	100	

[1] These companies include financial organisations where turnover might be misleading and those companies who completed the questionnaires anonymously.

(ii) The proportion of companies with no computer facilities available for planning is approximately the same for all but the highest turnover groups, which were always found to have some form of facility.

This appears to be consistent with the picture which emerges throughout the survey of the largest companies placing more reliance on planning and quantitative methods.

Few planning departments appear to rely solely on conversational systems for their computing, a mix of conversational and batch processing being very common. Indeed 75% of the companies had one or more forms of batch processing compared to the 47% who had a conversational system available.

Most of the survey companies have computer facilities available for planning but what, if any, models do they have and how often do they use them? In order to answer this question, the following broad categories of models were identified: forecasting, financial, corporate, marketing, personnel, production planning and 'one-off' models. Interest was restricted to 'computer-based' models so as to eliminate relatively simple *ad-hoc* models. As Table 3 shows, 60% of the companies had what they classified as a financial model; in fact this represents all but four of the companies with models of any kind at all. 38% of the companies claimed to have corporate models but with the present day trends towards simple corporate models it is likely that some confusion may have existed and that a great deal of overlap may exist between corporate models and financial models. We found that those companies with corporate models are a subset of those companies with financial models. Thus some 60% of the survey companies either had a corporate model, a financial model or both. Forecasting

Vol. 5
1977

TABLE 3. TYPES OF MODELS USED

Turnover groups £M / Model category	15–20 (%)	20–100 (%)	100–300 (%)	>300 (%)	Other companies (%)	Totals (%)
Forecasting	5	9	5	16	11	46
Financial	7	16	9	14	14	60
Corporate	7	9	2	11	9	38
Marketing	2	9	2	16	2	31
Personnel	0	2	5	5	0	12
Production	2	9	5	9	2	27
'One-off'	0	5	9	11	5	30
None of above models	7	11	7	2	4	31
% of total represented by each group	14	30	16	22	18	100

models were the next most popular and were claimed by 46% of the companies. Marketing models proved to be surprisingly common, particularly in view of the difficulties involved; so too did personnel models.

41% of those companies for which a production planning model is appropriate have one. (Some 30% of the companies do not produce a physical product and a further 5% are involved in the building and construction industry in which production planning models are not particularly applicable). The use of one-off models appears to be related to turnover in that their level of employment is relatively greater among the bigger companies: this is perhaps accounted for by a large and active OR group which is unlikely to be present in smaller companies. It is also noticeable that financial models are the most common in all the company groups with the exception of the highest turnover group where marketing and forecasting models are the commonest categories. Whether or not this is the reflection of a greater degree of marketing orientation it is impossible to say.

The relationship between the use of models and company turnover is further illustrated in Table 4 which shows the number of the above categories of models in use by the company against turnover. If the columns of this table are compared it can be seen that there is essentially little difference in character between the first three columns but in the fourth column, representing the high turnover companies, there is a visible tendency to use more of the categories of models. It would, therefore, appear that the larger and presumably more complex companies tend to make more use of models. We found a positive correlation of $r = 0{\cdot}47$ (significant at the 0·01% level) between company turnover and the number of modelling categories employed. It is not clear whether this is so primarily because the large companies have the resources necessary for model-

TABLE 4. NUMBER OF CATEGORIES OF MODELS IN USE AGAINST TURNOVER GROUPS

No. of categories of models in use \ Company turnover £M	15–20 No. of Co's	20–100 No. of Co's	100–300 No. of Co's	>300 No. of Co's
0	3	5	3	1
1	0	2	0	2
2	0	1	1	1
3	2	1	1	1
4	1	3	1	0
5	0	0	0	2
6	0	0	0	2
7	0	1	1	1
Totals	6	13	7	10

ling or because their organizational complexity compels the more extensive use of modelling techniques. There is, however, evidence to suggest that large companies not only tend to possess more models but also tend to use them more frequently than smaller companies. The frequency of use was found to be positively correlated with company turnover for financial, marketing and corporate models with correlation coefficients of $r = 0{\cdot}34$(significant at 0·1% level), $r = 0{\cdot}40$(significant at 0·1% level) and $r = 0{\cdot}69$(significant at 0·01% level), respectively.

No relationship was found to exist between frequency of use of forecasting models and turnover and the frequency of use of production planning models appears to be independent of company turnover.

The frequency of use of models was found to vary greatly from company to company but above a turnover of about £150M (the median turnover of the modelling companies) models tended to be used more frequently than for companies below this turnover. Generally speaking the high turnover companies used their forecasting, financial and corporate models on a weekly basis; amongst the lower turnover companies, quarterly or twice yearly use was most common. Exceptions did exist and one company was found to use its financial model on a daily basis while another company used its forecasting models several times daily.

Vol. 5
1977

As Table 5 shows, the high turnover companies tended to rate the impact of computers on planning as being more significant than the lower turnover companies which would seem to support the view that the larger organizations find the computer most useful.

Table 6 illustrates the response to the same question as Table 5 but this time the response is divided between those companies who have financial models (most of the companies who have a planning model of any sort have a financial model) and those who do not have a financial model. It is clear that those companies who have models tend to rate the impact of the computer on planning more highly than those who do not. It is noticeable that several of the non-modelling companies still rate the impact of the computer as significant. It

TABLE 5. ASSESSMENT OF THE IMPACT OF COMPUTERS ON PLANNING 1

Impact \ Turnover groups £M	15–20 (%)	20–100 (%)	100–300 (%)	>300 (%)	Other companies (%)	Total (%)
Very significant	3	0	8	5	2	18
Significant	0	17	3	17	13	50
Not very significant	10	10	7	0	5	32
Totals	13	27	18	22	20	100

TABLE 6. ASSESSMENT OF THE IMPACT OF COMPUTERS ON PLANNING 2

Impact	Modelling companies (%)	Non-modelling companies (%)	Totals
Very significant	15	3	18
Significant	38	12	50
Not very significant	12	20	32
Totals	65	35	100

would appear, therefore, that companies who have gone in for modelling find that they make a significant contribution to the planning effort, particularly in the larger companies. A disturbingly large number of companies did, however, feel the impact of computers to be 'not very significant'.

THE MODELLING PAY-OFF

There is often no easy answer to the question of modelling costs and pay-offs but it is frequently claimed that computer-based models enable plans to be made and, perhaps more importantly, revised more quickly; that they allow more detailed plans to be prepared through the inclusion of more alternative courses of action and through deeper analysis, all made possible by increased speed; and that overall more reliable plans result through the greater attention to detail made possible. In order to examine the extent to which these advantages had been felt, the survey companies were asked how far they would agree with the statements that computers:

1. Enable plans to be made/revised more quickly.
2. Enable plans to be made in more detail.
3. Result in more reliable plans.

As Tables 7, 8 and 9 show, the majority of companies felt that there were advantages to be gained from computerised planning; even those companies

TABLE 7. RESPONSE TO STATEMENT 1—COMPUTERS ENABLE PLANS TO BE MADE/REVISED MORE QUICKLY

Speed	Modelling companies (%)	Non-modelling companies (%)	Total
Agree strongly	29	3	32
Agree	29	24	53
Disagree	5	10	15
Disagree strongly	0	0	0
Total	63	37	100

TABLE 8. RESPONSE TO STATEMENT 2—COMPUTERS ENABLE PLANS TO BE MADE IN MORE DETAIL

Detail	Modelling companies (%)	Non-modelling companies (%)	Total
Agree strongly	27	7	34
Agree	29	22	51
Disagree	5	10	15
Disagree strongly	0	0	0
Total	61	39	100

without models, in general felt that there were advantages of speed and detail to be gained. In fact the survey companies were fairly unanimous that the use of computer-based techniques led to greater speed and detail but they were more evenly split on the question of reliability of the resulting plans.

While the survey companies as a whole believe the computer and the techniques dependent upon it have made a significant impact on planning, they do not believe that computers are indispensable to planning. When asked if they believed they could plan effectively without computer-based aids, the majority (92%) appeared to think so, as Table 10 shows. This is perhaps not surprising as many companies would argue that the discipline of planning and the thought processes it evokes are often more valuable than speed and greater detail in the plans themselves. As Argenti says [3] "it is a thousand times more important to apply the basic fundamental principles of corporate planning systematically than it is to be able to manipulate the tools of its technology". Those companies who have not adopted modelling techniques were as a whole positive that planning could be performed effectively without computers while those companies with models tended to agree but expressed reservations. The one company that was positive in its belief that computers were essential was a financial organisation and was, therefore, a largely atypical business. Thus we have the picture that many companies, particularly large companies, have and use models, but while they find them useful they do not appear to find them essential.

Vol. 5
1977

TABLE 9. RESPONSE TO STATEMENT 3—COMPUTERS RESULT IN MORE RELIABLE PLANS

Reliability	Modelling companies (%)	Non-modelling companies (%)	Total
Agree strongly	8	2	10
Agree	44	7	51
Disagree	13	23	36
Disagree strongly	0	3	3
Total	65	35	100

TABLE 10. IS IT POSSIBLE TO PLAN EFFECTIVELY WITHOUT COMPUTER-BASED AIDS?

Response	Modelling companies (%)	Non-modelling companies (%)	Total
Definite yes	7	26	33
Yes with reservations	46	13	59
No with reservations	3	2	5
Definite no	3	0	3
Total	59	41	100

FORECASTING TECHNIQUES

As we have already stated, forecasting models are second only in importance to financial models in terms of the proportion of the survey companies that employ them. The large companies in particular seem to place even greater emphasis on forecasting models than they do on financial models. This is perhaps not surprising since companies which do not engage in some form of forecasting must be very rare indeed, while not every company by any means is interested in detailed financial planning. Forecasting in one form or other is essential to every business but not every company uses a detailed model to make these projections. In fact as Table 11 shows, subjective methods and statistical extrapolation using simple statistical methods or even curve fitting 'by eye' are the most common methods of forecasting.

Although subjective methods are very popular, only 15% of the companies rely on this technique exclusively while 37% of the companies use only one of the four techniques. It is interesting to note that one of the very large companies did in fact rely exclusively on subjective methods, a practice one would expect to be confined to only the smaller companies.

CORPORATE MODELS

In 1973, Grinyer and Wooller [1] conducted a random survey of "The Times 1000 Companies" and found that some 9% of the companies had a corporate model or were developing one. Since then the growth rate in corporate modelling has been considerable and Grinyer [4] suggests that over a thousand companies now either have or are developing a corporate model. As stated earlier in this paper, 38% of the survey companies were found to have a corporate model and 60% were found to have either a financial model or a corporate model. A further 13% of the companies were found to be developing a corporate model. It would, therefore, appear that the number of large UK companies engaged in

TABLE 11. METHODS OF FORECASTING USED

Method	%
Subjective estimates	73
Statistical extrapolation	76
Operational research or econometric models	44
Technological forecasting	29

corporate modelling has increased significantly between 1973 and 1976 even allowing for possible bias in our respondents. In addition to this, a further 16% of the companies said they would like to develop a corporate model, indeed all but one actually had plans to develop a corporate model, although it should be noted that most of these companies already had what they described as a financial model.

The numbers of corporate modellers are not only increasing rapidly but they are also a resolute species. Of the sample companies, 27% had at some time developed a corporate model which subsequently fell into disuse; not discouraged, two-thirds of these companies have since developed a further corporate model which is in use today and the remaining third currently operate a financial model. No instances were found of companies giving up the effort completely.

The reasons given for the failure of the model were in general the classic reasons given in the literature. Of the eleven companies who had built a corporate model which has since been abandoned, the most common reason, given by five, was difficulty in collecting the necessary data. Three companies found their models too complex to be useful, while four cited unreliable results as a cause. Behavioural problems also arose: two companies cited lack of senior management support; three lack of staff acceptance; one lack of training and one the departure of the senior management sponsor of the model. In a further case, the model was invalidated by a change in the company structure.

In addition to these companies, one large organisation, with a well-known and respected OR group, gave up in its attempt to build a corporate model as it found it too complex and inflexible to handle the range of possible company futures and has not made a second attempt at modelling since then. A second company was persuaded to build a corporate model by management consultants. This model represented the activities of one division and was composed of a suite of programs built by the consultants in conjunction with their own OR group. The resulting model was highly detailed and consequently complex with a massive data input and a vast range of outputs. Indeed the model was so massive that it was just too large to handle and only one full run was ever made, although some of the sub-units of the model were used more extensively. In contrast to this approach, this particular company has now, in line with the

present trend, developed a simple, highly aggregated model of its UK operations which it feels is more appropriate to its information needs and is considering applying the technique to financial models at divisional level.

Many organisations seem to find the development of a corporate model worth-while as illustrated in Fig. 1. It seems worthy of note that only two of the twenty-one companies with either a corporate or a financial model judged corporate models to be of limited value. Clearly there is a marked tendency for those companies with corporate (or financial) models to value them more highly than those without. This would suggest that many of the companies without models might in fact find them of value if they made the developmental effort.

VIEWS OF COMPANIES WITHOUT CORPORATE MODELS

We found that 51% of the survey companies have, or are developing, a corporate model and that a further 16% have plans to develop one in the future. Of those companies without a corporate model 29% would like to develop one, 50% might like to and 21% definitely would not. Overall, therefore, most companies were quite favourable towards the idea of a corporate model. Why then have they not developed one? Table 12 gives a summary of the reasons given by these companies.

As Table 12 shows a wide variety of reasons were given for not having developed a corporate model and it is difficult to fit an overall pattern to the response. Two factors are, however, clearly apparent as being more common than the rest: insufficient personnel for the task and the belief that a corporate

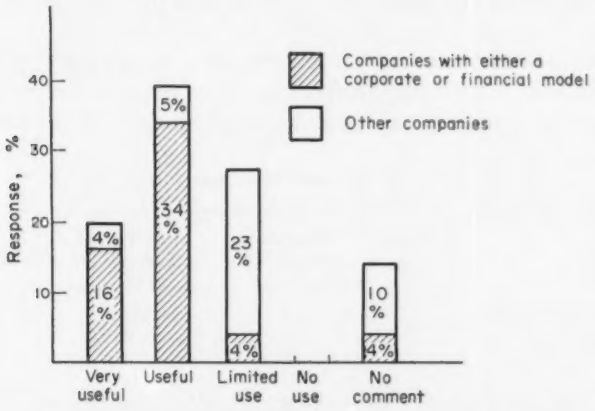


FIG. 1. *Usefulness of corporate models as aids to management.*

TABLE 12. REASONS GIVEN FOR NOT HAVING DEVELOPED A CORPORATE MODEL

Reason given	% Response
Corporate models felt to be too inflexible	4
Doubt benefit justifies the cost	8
Insufficient computer facilities	4
Insufficient personnel	33
Lack necessary expertise	8
Inadequate input data	8
Lack of senior management sponsor	8
Not appropriate to type of business	25
New to coporate planning	8
Do not require necessary detail to justify model	4
Expect reorganization of group	4
Cost involved	4

model is not appropriate either to the business or the structure of the organization. In general, this latter reason was given by those companies who had no intention of developing a corporate model though one of this group of companies felt corporate models were too inflexible for its needs and another felt that such detailed information was not appropriate to the company. It is noticeable, however, that about half the companies cite as their reason a lack of resources whether of manpower, expertise or computer facilities and as such these resources could be bought in. It is perhaps of some concern that so many companies felt a corporate model was not appropriate to their business. Certainly a model of some sort, even a simple financial model, could be of value in principle to almost any organisation. The value of a model is determined more by the approach to planning than by the nature of the organisation.

CORPORATE MODEL USERS AND BUILDERS

As Table 13 shows, accountants are the most common group of users of corporate models in the survey companies. However, it is interesting to note that company directors use the corporate model in 59% of the companies, a

TABLE 13. THE PRINCIPAL MODEL USERS

Use by	Disused model (%)	Operational models (%)	Model under development (%)
Operational researchers	13	18	43
Computer specialists	13	0	14
Accountants	50	71	71
Planners	63	53	57
Directors	25	59	14

surprisingly high figure, while the figure of 53% for planners is lower than was expected. This could be explained on the basis that in the smaller companies of the sample, directors themselves carried out work done by planners in the larger companies.

It is noticeable that the present generation of corporate models is used extensively by directors while the now disused models were used by directors in only a limited number of companies. This is perhaps testimony to both the increased awareness of top management to the potential of corporate models and the trend towards more simple models. Indeed it was found that some 39% of chief executives had actually used a corporate model, although it seems more usual (even where an on-line model exists) that the chief executive instructs someone to use the model on his behalf rather than use it personally. The use of corporate models by operational researchers is lower than was expected except at the development stage where it is to be anticipated that OR will feature more extensively.

The respondents were asked to estimate the relative contributions (as percentages) of the various professions to the building of the company's corporate models. The work of Grinyer and Wooller [1] would indicate that the importance of the contribution of operational researchers and computer specialists should be less in the operational models than in the disused models, since they report a trend towards models of greater simplicity. In fact as Table 14 indicates, there appears to be little significant difference between these two groups, although our general experience does suggest some increases in the parts played by accountants and planners.

It is of interest to note that of the 23 models studied, only three involved consultants and only one model was found to be of the 'ready-made' type and this company had a second, internally-developed model as well.

This particular survey has failed to detect a trend towards more simple models from the disused models to the models operational at present, although this is based purely on whether the researchers judge their models to be 'top-down', 'middle-up' or 'bottom-up'. Table 15 summarises the results and shows

TABLE 14. THE RELATIVE CONTRIBUTION OF THE MODEL-BUILDERS

Relative contribution by	Disused model (%)	Operational models (%)	Model under development (%)
Operational researcher	17	15	14
Computer specialists	10	28	10
Accountants	39	44	29
Planners	29	34	39
Consultants	6	5	10
Directors	0	5	0

TABLE 15. CORPORATE MODEL LOGIC-STRUCTURE

Type	Disused model (%)	Operational models (%)	Model under development (%)
Top-down	43	38	33
Bottom-up	14	12	17
Middle-up	43	50	50

that 14% of the disused models and 12% of the operational models were 'bottom up'. As 17% of the models under development were said to be 'bottom-up' there is no evidence here to suggest a reaction against detailed models.

Finally, Table 16 summarises the uses to which the models are put. The preparation of financial projections appears to be the most common use of corporate models, followed by long-range planning and then the evaluation of policy alternatives. Use of corporate models appears to be less common in short-range planning, the preparation of financial reports and in decision analysis.

CONCLUSION

Large numbers of companies appear to be using modelling techniques as part of the planning and strategic management processes. In general, there appears to be a fairly even spread of the technique across all company sizes except the very biggest, who appear to have a wider variety of models available and also to use them with greater frequency than the smaller companies. These larger companies also appear to value the contribution of models more highly than the smaller companies. While most companies found the use of computers in their planning process enabled plans to be made and revised more quickly and greater detail to be incorporated in them, they felt that computers were not essential to effective planning.

TABLE 16. USES OF CORPORATE MODELS

	Disused model (%)	Operational models (%)	Model under development (%)
Evaluation of policy alternatives	50	63	100
Financial projections	83	88	85
Analysis for decision making	0	44	57
Long-range planning	67	81	85
Short-range planning	0	38	28
Preparation of financial reports	17	44	14

In the more narrow area of corporate modelling, there appears to have been tremendous expansion since 1973. These models are often used by managers surprisingly high in the organization's hierarchy, over half of the companies stating that directors were amongst the principal users of the model. Many other companies appear to be interested in corporate modelling but are deterred by lack of the necessary resources. Other companies, however, exist who feel models are not appropriate to their types of business or their structures and we would suggest that such companies should be less sceptical.

REFERENCES

1. GRINYER PH and WOOLLER J (1975) Computer models for corporate planning. *Long Range Planning* **8**(1) 14–25.
2. NAYLOR TH and JEFFRESS C (1975) Corporate simulation models: A survey. *Simulation* **24**(6), 171–176.
3. ARGENTI J (1974) *Systematic Corporate Planning*. Nelson, London.
4. GRINYER PH (1976) *Corporate Modelling*. Data for Decision Conference, London, May 1976 (Online Conferences Ltd.)

ADDRESS FOR CORRESPONDENCE: *Professor JC Higgins, Management Centre, University of Bradford, Emm Lane, Bradford, West Yorkshire BD9 4JL*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977. Pergamon Press. Printed in Great Britain

A Planning Model for Underground Mines—An Application in a Developing Country

BERNARD A LIETAER

University of Louvain

(*Received October 1976*)

This paper describes the planning of an underground mine using linear programming. The objective of the plan was to determine the optimal schedule of mining works to be performed in order to maximise the contribution to the company over a specified time horizon. Details of these mining works, and also the decision variables used in the model are included in the paper. The model is now being used operationally in the Casapalca mine, and is being extended to the other mines of the company.

1. 5
77

I. INTRODUCTION

THIS paper documents the development of a new approach to model underground mine planning. A number of models exist for open pit mine optimization with particular emphasis on either open pit design or equipment assignment (See for example [1, 2, 3, 4]) but optimization models for underground mines, to our knowledge, had not been designed so far.

The idea to attempt the design of an optimization model for underground mines came as an extension of the successful development and implementation of two LP models for the smelting and refining operations of Centromin-Peru (previously Cerro de Pasco Corporation) respectively designed in 1971 and 1972.

Although the mining model presented in this paper was conceived to meet a specific set of requirements particular to Centromin, it is adaptable to the majority of underground mine operations in the world.

II. BACKGROUND

Centromin-Peru is the largest integrated mining, smelting and refinery company in Peru. Although total output of over US $300 million places Centromin only in the medium size league for non-ferrous metals companies, it is definitely one of the most complex in the world as it simultaneously produces

no less than 16 end products of which major ones include copper, zinc, silver, lead, cadmium, antimonium, gold, bismuth, tungsten, tellurium and indium.

Its relative importance in the Peruvian economy gives Centromin and its modeling experience an additional dimension. About 50% of hard currency earnings in Peru come from the mining sector, and Centromin is responsible for about half of this amount. After the 1974 nationalization of the company, all the models were tested for both profit maximization (from a company viewpoint) and foreign exchange optimization (from a country viewpoint). Until now, both optimizations have yielded similar results, and no conflicts have developed between both objectives.

Figure 1 presents a schematic outline of the Centromin operations system. It includes six mining zones and concentrating plants scattered in the higher ranges of the Andes of Central Peru, a smelting and refining complex in La Oroya, two railroads, and a network of four hydroelectric plants.

The original idea, proposed in 1970, to develop one single large optimization model for the entire system, proved impractical because of the size and com-

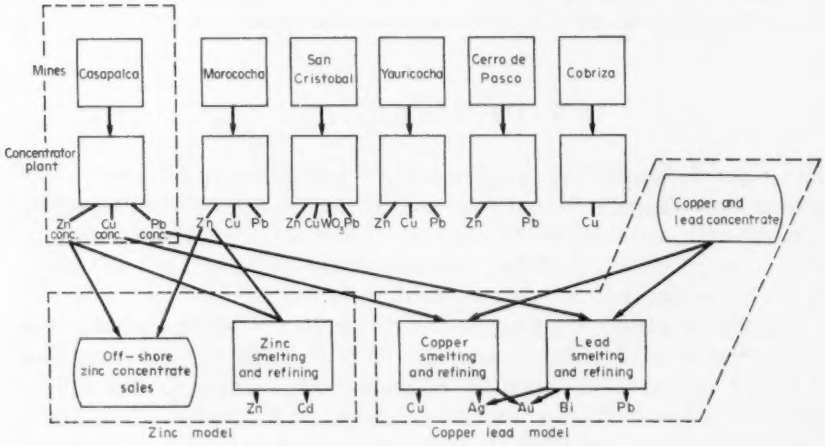


FIG. 1. *Centromin operations system.*

plexity of the resulting matrix (the mining model discussed in this paper, corresponding to one single mine, turned out to require several thousands of variables). Therefore a set of submodels were developed with interconnection possibilities (providing as input of one model the output of another). Although this break-up can result in theory in an overall sub-optimization, pragmatic computing limitations and desirable user flexibility were dominating considerations.

The sub-models which are already operational in this system are the zinc model and an integrater copper–lead model (see Fig. 1) which optimizes the total system using as input the concentrates.

As Centromin produces zinc concentrates in excess of its smelting capacity, the zinc model, designed in 1970 and implemented in 1971, optimizes the allocation of the zinc concentrates between the company smelting plant and a series of off-shore concentrate sale contracts. After the nationalization of the company and of all off-shore sales of metals and concentrates in Peru, the model has been expanded to include all zinc concentrates produced and all the concentrate sales contracts in the country. This application previously documented has been credited with improving dollar proceeds of Peruvian zinc concentrates by several $ Million per year.[1]

As Centromin's own concentrating plants produce only between 45 and 60% of the copper and lead concentrates required for its Oroya smelters, the copper and lead model have the objective to optimize the purchase of other non-company concentrates in order to 'top off' the metallurgical balance requirements (some fifteen metallurgical relations have to be met) for the smelter beds.

After the successful implementation of these two models, a set of 'desirable metallurgical characteristics' became identifiable for the company's own concentrates, taking into account the impact of each concentrate on the total metallurgical balance of the beds.

The next logical step became then to attempt to organize the mining plans in each of the company's mines to produce a mineral mixture which would allow to the concentrating plants to adapt their production to these metallurgical requirements. Hence the need for an optimization model for the company's mines.

The capabilities to meet this need have also been built up over this time period.

The first two models had allowed the in-depth familiarization of a group of systems analyst and engineers with LP models. A 'model department' was set up in the Systems Division of Centromin, which trained seven specialists in model design, usage and interpretation.

After several years of successful experience, it was decided to complete the modeling of the full scope of the company's operations by attempting to include the mining and concentrating activities.

The design of this new planning model was initiated in November 1974 and was performed by a team including at various times between 7 and 10 mining, industrial and systems engineers.[2]

[1] A documentation of this model was contributed by the author [5]. Also see *Business Week*, 26 August, 1972 'A computer model to upgrade zinc profits'.

[2] Key team members include M. Moya, D. Sarmiento, J. González, V. Argüelles, A. Odria, V. Quispe, A. Chang, all from Centromin, under the technical guidance of the author retained as external consultant.

The entire effort, including design, systems programming and testing, the industrial engineering work to evaluate equipment standards, and implementation amounts to about 12 man-years.

The full system was tested on operational data during early 1976 and entered in routine managerial usage in May 1976.

III. DESCRIPTION OF THE REALITY TO BE MODELED

Five out of the six mines of Centromin are underground operations. Therefore, the pilot mine which was selected to attempt the modeling effort was also underground. The Casapalca copper, lead, zinc and silver mine located 120 km from Lima at 4200 m altitude on the Western flank of the Andes, was selected for the pilot modeling work.

The scope of the model includes the mine and the concentrator plant which together are referred to as the Casapalca camp.

The mining activity consists of extracting 1850 tons/day of copper, lead, zinc, and silver sulfur complex ores from some 40 simultaneously active stopes from some 30 levels at altitudes ranging between 4850 and 3640 m in altitude. The deposit is geologically very diversified as it consists of 27 interconnecting vertical veins.

Vol. 5
1977

The predominant mining technique in Casapalca is 'Shrinkage', which involves four successive phases of mining work:

- *Development:* is the phase of the mining work necessary to gain access to the mineral to be extracted. During this phase, since the drifts are normally driven on a vein the material is usually rich enough to be sent to the concentrator. On the average in Casapalca, the development phase of a typical stope last 2–3 months.
- *Preparation:* is the phase when the lateral raises, the box holes and the first sublevel needed to start systematic breakage are driven. The material produced during the phase is rich enough for concentrator feed and is used accordingly. The average length of time for preparing a stope in Casapalca is 2 months.
- *Breaking:* when the stope is being broken, after each cut, about one third of the mineral is being extracted in order to lower the level at the floor sufficiently to start a new cut. Average time in this phase is 5–8 months for the entire stope.
- *Drawing:* after the stope is completely broken, the mineral which remains accumulated (about 2/3 of the original broken tonnage) is drawn from the stope according to feed requirements in a time ranging between 1 and 6 months.

The operations within the mine have a large number of technical capacity constraints and a systematic survey of all possible limitations determined the critical factors which limit total production of the mine. These are: manpower, drilling equipment (stopers and jack-legs) mining cars, and hoisting capacity.

The mineral extracted from the mine is fed to the concentrator where it is successively crushed, ground, and separated by flotation in three different concentrates (copper, lead and zinc concentrates with respectively typical metal content of 30% in copper, 54% in lead, and 60% in zinc). The limiting factors in the concentrating factors in the concentrating plant in Casapalca were determined to be the grinding (1850 tons of mineral/day) and the flotation cells (3620 tons/month; 405 tons/month; 2156 tons/month respectively for zinc, copper and lead concentrates).

These are the general outlines of the reality to be modeled. In the balance of this paper, the approach to modeling these operations is described first in its general structure, and then by defining the specific variables and equations. Finally some practical applications are documented.

IV. GENERAL STRUCTURE OF THE MINING MODEL

The objective of the underground mining model is the optimization of 'mining plans', i.e. the selection and scheduling of the mining works to be performed in order to produce a mineral feed to the concentrator that represents a maximum contribution to the company.

In more practical terms, this objective was expressed as preparing a combination of mining works which fulfill the following four simultaneous requisites:

- Allow the production of concentrates with the highest possible value.
- At the lowest possible total variable costs for mining and concentrating (as usual, as fixed costs are not affected by the decisions, they are not considered in the model).
- Taking into account all the operational restrictions of the mine and concentrator (equipment available, capacity constraints, etc.).
- And reflecting the company's general policies (keeping a constant level of mineral reserves in the mine, etc.).

The structure of the model becomes then a mix optimization LP model with two significant variations: it is a two-tier mix model, and it is multitemporal. Both aspects require some discussion.

(a) *Two-tier mix model*

As Fig. 2 illustrates, the selection of the stopes for the mining plan have to fulfill a first set of mine operations constraints, producing a blend for the total mine. A second level of optimization occurs at the concentrator level where one or several of the concentrator operations options are selected.

(b) *Multitemporal model*

Mining is a continuous process, therefore it is necessary to develop and prepare stopes continuously in order to replace the broken and extracted ore. Consequently, an 'optimal plan' for this month includes also development works which will produce mineral only in future months.

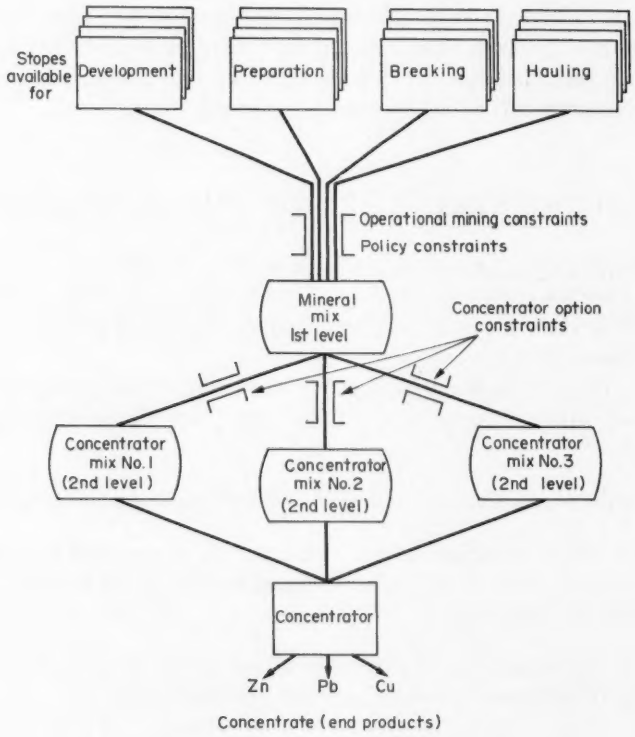


FIG. 2. *Two level mix problem of the mining model.*

Hence the model has to allow the optimization of a successive series of time periods, each one depending on the works performed in previous periods. The 'optimal mining plan' to be determined is the one which optimizes the entire

Vol. 5
1977

planning horizon, and not only a plan optimal for the first month or semester. Therefore the model breaks up the planning horizon in a series of planning periods (for example a 1 year plan broken up in four successive quarterly planning periods).

Ideally, the optimal plan should take into account the chain effect of this planning process over the entire life time of the mine. However, data such as prices, geological reserves, and operational characteristics, is not reliable as one projects further into the future, and a pragmatic and shorter planning horizon was accepted (from 1 to 10 years depending on the decisions for which the model is run. Accordingly, the model has the flexibility to optimize simultaneously up to 10 periods, each of a length varying from 1 month, to 1 quarter, or 1 year).

However, in order to approximate the concept that this planning horizon is not the true horizon of the problem an additional period was added to the planning chain, a special 'residual period'. This 'residual period' is not part of the actual mining plan, but it has as objective an evaluation of the impact of any work started but not finished at the horizon of the problem.

V. DECISION VARIABLES AND MODEL EQUATIONS

Only the most important variables and equations of the model are described here, in order to provide a general idea of the model's matrix.

(a) *Decision variables*

The decision variables of the model are tons of minerals of a particular stope, at a given mining phase, destined to feed one of the concentrator options, at a given time.

Therefore the decision variables are called:

i e k t: where *i* name of stope (up to 300 different ones);
e one of the four mining phases: development, preparation, breaking, drawing;
k one of 3 different concentrator operational options;
t time period, from 1 to T (up to 10) for planning periods within the time horizon, and $T+1$ for the residual period.

(b) *Other key variables*

Other key variables of the model include:

- Mineral production variables, which sum the total production by mine level, and the total for the mine.
- Feed mix variables, which sum the metal content of the production for each metal, the mineralogical composition for each mineral and the impurities which affect the flotation process.

- Concentrate production variables, which relate the mine production to the end products of the concentrator plant (zinc, lead and copper concentrates) through the recoveries, corresponding to each option of the concentrator.

- Cost variables, identifying variable costs for each mining phase in each of the planning periods.

- Income variables, accumulating the economic value of the concentrates produced.

- Equipment variables, accumulating the total man-shifts, stopers-shifts, jacklegs shifts, hauling-hours, and mining car-shifts used for the mining plan.

- Policy variables, particularly reserves of minerals ready for development, preparation, breaking and drawing.

(c) *Objective function*

Maximize Gross Contribution Z=(Value of concentrates produced)
−(Variable Mining and Concentrating Costs.)

The value of the concentrates produced is determined through the usage of the other LP models of Centromin, where concentrates are used as input (see Fig. 1).

For example, the different types of zinc concentrates which correspond to the various operational options of the concentrating plant, are first evaluated and optimally allocated with the zinc optimizing model.

The zinc model determines on an overall optimization basis for all zinc concentrates of Peru that this grade should not be processed in Centromin's own smelter in Oroya, but sold off-shore to a specific sales contract (for example to a Japanese or a US smelter). The conditions of this sales contract and its valuation clauses provide the data needed to assign the correct price to the various grades of zinc concentrate of Casapalca.

The mining model then determines which of these options is optimally produced in the Casapalca camp. It is theoretically possible that several iterations between both models might be necessary to obtain a true overall optimal solution, however, this has not turned out to be a problem in actual case.

(d) *Other model equations*

Among the key equations of the model should be mentioned:

- Capacity constraints on tonnages for each stope, and total concentrator tonnage capacity both for grinding of minerals (as raw material input) and for flotation of concentrates (as end product).

- Concentrator feed constraints determining acceptable minima and maxima for each concentrator option, on assays for metals, mineralogical composition and impurities.

- Manpower and equipment capacity constraints for each planning period, limiting total requirements in manpower, stopers, jacklegs, hoisting and cars to available maxima.

- Policy constraints through a set of minima on mineral reserves of different types for each planning period.

- Sequence constraints, a set of equations which ensure that the sequence of the different phases of the mining method are followed, specifically insuring that a stope has been broken in previous time periods before it can be hauled, has been prepared before being broken, and has been developed before being prepared.

This last set of equations is in fact only an approximation of the correct modeling approach which would require mixed integer programming (because for example the development phase should be 100% complete before preparation can begin). However, mixed integer programming algorithms on several thousands of variables cannot yet be considered practical in terms of computer solution time.

VI. THE MINING MODEL SYSTEM

Many different types of analysis are required for the mining model. The data preparation requirement for any given run, being very large, makes the model in fact 'tailor-made' for each run, by the usage of a set of programs and computer based data banks.

The full mining model system (Fig. 3) includes a data bank, a matrix generator, an MPSX package and a report generator.

The 'Casapalca data bank' covers all the geological, operational and industrial engineering data needed for all the mining zones to be considered in Casapalca. For instance, for any given stope, the data bank includes vein metal assays and metallurgical composition, as well as manpower and equipment standards, the present status and previous work performed, rehabilitations costs and time if applicable, etc.

This data bank has an updating program, a report production program to prepare managerial reports about the status of the Casapalca Mine, as well as a data selector program which extracts from the data bank what would be required for any particular model run.

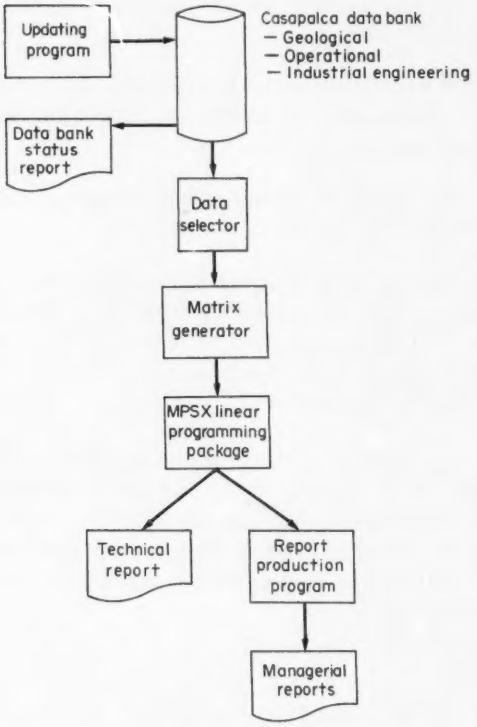


FIG. 3. *The mining model system.*

The matrix for the LP package is automatically generated to the specifications of a particular run, as there is a flexibility in terms of number and types of stopes to be considered for the plan, the total planning horizon and the size of each planning period (for example a 1 year plan with quarterly periods, or a 10 year plan with yearly periods).

A 'normal' run at Casapalca uses quarterly data and has a matrix size varying between 2500 variables and 750 equations and about 4000 variables and 2000 equations.

Average computer time to reach optimality in the IBM 370/135 of Centromin is about 1/2 hour for the lower end of this scale, and up to 5 hours for the upper end.

Besides the normal technical output from the MPSX package a special managerial report is produced, which presents the results in a format directly usable by the Mine and Concentrator Superintendents.

VII. CONCLUSIONS

As compared to the zinc and copper–lead models which have respectively already 4 and 2 years of successful usage, the mining model has been operationally used only since May 1976. The first practical applications allowed to produce a mining plan which stabilized significantly the head assays to the concentrator plant, as well as a more efficient usage of manpower and equipment (particularly improved scheduling mix of stopers and jacklegs). Preliminary economic impact valuations amount to about $1 Million for a 6 month period.

The main benefits which seem attributable to the model, besides the economic improvements and better equipment utilization, is the possibility to explore the feasibility of a wide range of concentrator options and verify their impact on the mining plan. Coordination between requirements of a concentrator plant and an underground mine has always been a tricky task, and the formal quantification of all critical variables makes the gradual and controlled experimentation of different mixes of minerals to improve concentrator recoveries an easier task.

The same model structure has been adapted to a variety of other underground mining techniques in particular hydraulic cut and fill, arch-back, and square sets. A program has been initiated to generalize the Casapalca pilot model to the other mines of the company.

The same model, adapted to specific local conditions, should also be able to improve an underground mine planning in other companies, introducing OR in an area where few modeling efforts have been made until now.

REFERENCES

1. HEWLETT RF and REDMON DE (1963) *Design and Application of some Mathematical Models for Mine Systems Analysis*. US Dept of Interior, Bureau of Mines.
2. CARLSON TR, ERICKSON JD, O'BRIAN DT and PANA MT (1966) Computer techniques in mine planning. *Mining Engng* **18** (5).
3. REDMON DE (1964) *Determining Mine-Production Schedules by Linear Programming*. US Dept of Interior, Bureau of Mines.
4. JOHNSON TB (1969) *Improving Return from Mine Products through Use of OR Techniques*. US Dept of Interior, Bureau of Mines.
5. SALKIN HM and SAHA J (1975) *Studies in Linear Programming*. North Holland, Amsterdam.

ADDRESS FOR CORRESPONDENCE: *Bernard A Lietaer, 19, Avenue Robert Dalechamp, 1200 Brussels, Belgium.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977. Pergamon Press. Printed in Great Britain

Structural Sources of Personnel Problems in Multinational Corporations: Third-Country Nationals

YORAM ZEIRA

Graduate School of Business, Tel-Aviv University

and

EHUD HARARI

Faculty of Social Sciences, The Hebrew University, Jerusalem

(Received July 1976; in revised form September 1976)

The paper examines the extent to which multinational corporations (MNCs) whose staffing policies approach the genuine multinational model are free of the dysfunctions found by the authors in earlier studies of ethnocentric MNCs. The findings are based on comparative research in 51 MNCs, and intensive diagnostic studies of seven subsidiaries of MNCs operating and headquartered in four continents. The findings may be summed up as follows: (1) certain, but very few, morale problems have indeed been eliminated; (2) others are still present; (3) others are still present and their effect has intensified; and (4) new problems have appeared which are unique to the staffing policy approximating to the genuine multinational model. These findings shed new light on several of the assumptions underlying this model. The general conclusion is that each prevalent type of staffing policy has several sources of morale problems—some inherent in its uniqueness, and others shared by the alternative staffing policies—irrespective of the personal qualifications of the managers. Therefore, it is unrealistic to expect that a change toward the genuine multinational model would solve the basic morale problems prevalent in the dominant type of MNCs.

THIS study is an advanced phase of an ongoing project focusing on personnel problems in multinational corporations (MNCs) of various structures. The ultimate goal is to reveal the unique personnel problems inherent in the prevalent staffing policies, and to analyze the functions and dysfunctions of each particular staffing policy for both the corporation and the employees.

The previous phase concentrated on ethnocentric [7, 12, 13] MNCs in which most key positions at headquarters (HQ) and in the subsidiaries were reserved for parent-country nationals (PCNs); this being the dominant structure of MNCs

The research reported in this article was supported by the Israel Institute of Business Research at Tel-Aviv University.

throughout the world [11, 17, 18], it was a natural point of departure. We analyzed the sources of morale problems in the light of problems reported about uni-national corporations, and singled out those sources we considered unique to MNCs. More detailed reports of that phase of the study and its findings appeared elsewhere [3, 4, 23, 24, 25, 26].

The policy of reserving key positions for PCNs has been recognized as problematical in three spheres: *international politics*—as being incongruent with host-country nationalism; *employee morale*—as frustrating host country nationals' (HCNs') personal ambitions for promotion, and falling short of fulfilling their needs and meeting their expectations [8, 9, 16, 17, 21, 22]; and *organizational effectiveness*—as perpetuating mismanagement of human resources and lacking compatibility between organizational practices and environmental characteristics [6, 8, 15].

Consequently, an alternate staffing policy has been advocated [2, 5, 6, 7, 10, 20, 27] for achieving genuine multinationalization: filling all key and top positions—both at HQ and in the subsidiaries—with the best qualified persons recruited from throughout the MNC system, irrespective of ethnicity or nationality. This would not only look right politically, it would also open new opportunities for competent individuals, and improve organizational effectiveness. In particular, such a policy would provide the following benefits: development of executives with international experience and international point of view; acquisition of a global approach to opportunity finding and decision making; perception of HQ and subsidiaries as integral part of one organic entity; creation of a multinational atmosphere; and initiation of organizational renewal projects in the subsidiaries based on expatriates' international experience.

Vol. 5
1977

It is doubtful whether any existing MNC fits the genuine multinational model described above. Several, however, have been consciously trying to approximate it. Moreover, as it has been gaining publicity and respectability among academicians and practitioners as a model for emulation, some even consider it to be the wave of the future.

Advocacy of this policy is based on five related assumptions:

(1) Highly competent employees are available not only at HQ, but also in the subsidiaries.

(2) International experience is a condition for success in top positions.

(3) Managers with high potential and ambition for promotion are constantly ready to be transferred from one country to another.

(4) The competent and mobile managers have an open disposition and high adaptability to different conditions in their various assignments.

(5) Those not blessed initially with an open disposition and high adaptability can acquire these qualities as their experience abroad accumulates.

On the basis of the literature, we hypothesized that those MNCs approaching the genuine multinational model would indeed be free of the dysfunctions revealed in our previous studies. For that purpose, we surveyed several such MNCs; we also conducted a diagnostic study in one MNC, focusing on a subsidiary headed by third country nationals (TCNs). This category of employee has been given special attention in our study because genuine multinationalization is reflected in their increasing placement in key and top positions.

We examined the extent to which morale problems typical of MNCs of the dominant type have been eliminated in MNCs with a more multinational structure. Our findings may be summed up as follows:
(1) certain, but very few, morale problems have indeed been eliminated; (2) others are still present; (3) others are still present, and their effect has intensified; and (4) new problems have appeared which are unique to the alternate structure.

THE STUDY

The previous phase of our study consisted of three stages: (1) A diagnostic study of employees at all hierarchical levels in six subsidiaries operating in France, Israel, Japan and the US; their HQ were located in four different countries in four continents. (The research reported here is a replication of that diagnostic study in an American MNC approximating the genuine multinational model.) (2) A questionnaire responded to by 51 personnel vice presidents of leading MNCs (petroleum, chemicals, cosmetics, automobiles, airlines, banks, international hotels, electronics). (3) Personal interviews with 39 of the respondents (or their deputies) and members of their staff at HQ. The respondents in all three stages belonged to three categories of employees: PCN, TCNs, and host-country nationals (HCNs).

The MNC reported on here is headquartered in the US and has over 60 subsidiaries abroad. It has been HQ policy to fill top positions in its subsidiaries with TCNs; when this has not been possible, with PCNs, but never with HCNs, "in order to avoid in-breeding and stagnation". While a few *key* positions at HQ have been held by TCNs, no TCN has ever held a *top* position at HQ; this practice has been deliberate, and in all likelihood will persist in the foreseeable future. Thus, approximation of the genuine multinational model has been in terms of the rotation of TCNs and the encouragement of outstanding HCN managers to assume top positions in third countries.

The subsidiary diagnosed had 612 employees: 605 HCNs and seven TCNs. The seven TCNs held all of the top positions. There were no PCNs in the subsidiary. A random sample of 43 was chosen from among the HCN employees, and the seven TCNs were then added. A questionnaire was administered to the sample; then 25 of the respondents were intensively interviewed: all seven TCNs, and 18 of their subordinates. The interview schedule was partially

structured and partially unstructured. By asking the respondents to elaborate on their responses to the questionnaire, it was possible to ascertain the validity of the instrument. In addition, information about morale problems not covered by the questionnaire was elicited. The findings reported below are based, therefore, on data obtained in the questionnaire and validated in the interviews; others—on interviews with HCNs and TCNs in the subsidiary diagnosed and interviews with TCNs in other MNCs which participated in previous phases of our study.

The questionnaire measured satisfaction with current policies and behavior and perceptions of the desired ones. The items were structured on a Likert-type scale ranging from 1 to 5; the higher the score, the higher the satisfaction or the intensity of the attitude. Scores of HCNs and TCNs were analyzed, and the differences between the scores of the two categories were tested (two-tail t test at the 0·05 level).

FINDINGS

Vol. 5
1977

We identified the following sources of morale problems as prevalent in MNCs of the dominant type: (1) blocked promotions; (2) transfer anxieties; (3) income gaps; (4) unfamiliarity and adaptability difficulties; (5) avoidance of long-range projects; (6) inappropriate leadership style; (7) non-participative decision making and screening of information. Below we discuss each source—summarizing its characteristics in the ethnocentric MNCs, and then reporting corresponding findings in the subsidiary headed by TCNs.

Promotion. In the MNCs of the dominant type, the policy of reserving most key positions at HQ and in the subsidiaries for PCNs meant blocked promotions for HCNs, irrespective of their competence, experience and loyalty. Naturally, they were dissatisfied. The PCNs[1] were much more [4] satisfied with their promotion opportunties than were the HCNs. Nevertheless, the PCNs were critical of three aspects of promotion in their MNCs: (1) absence of corporation-wide manpower planning; (2) absence of career ladders; and (3) promotion to key positions at HQ of PCNs who avoided serving abroad. The PCNs felt that rather than being rewarded for their achievements abroad, they were having their term of service extended beyond what they considered appropriate. In the meantime, those who stayed at HQ, although less qualified, were promoted. Moreover, those promoted at HQ did not know the PCNs well because of the latters' extended stay abroad, and were not familiar with their achievements in the subsidiaries.

[1] Whenever reference is made to differences between categories of employees stating higher or lower levels of satisfaction or intensity of attitude, the differences have been found to be statistically significant.

In the subsidiary which we diagnosed, the TCNs were as dissatisfied as the HCNs[2] (2·71 *vs.* 2·72) [4]. In particular, their dissatisfaction was expressed in the strongest terms about the extent to which the policy of blocked promotions was justified (1·58 *vs.* 2·00). In this respect, both HCNs and TCNs were more critical than their counterparts in the other subsidiaries headed by PCNs.

TCNs' dissatisfaction stemmed from two related factors which PCNs in the other corporations were spared. First, since top positions in the subsidiaries were reserved for non-HCNs, TCNs holding top positions found it impossible to return to the subsidiary in their own country; this would have involved a demotion in the organizational hierarchy. For PCNs, on the other hand, assignments abroad were designed to equip them with international experience —and more often than not, their return to HQ was followed by a promotion.

Second, each promotion for a TCN was ordinarily associated with another transfer to a subsidiary in a new third country. While the first or second transfer might have been enticing for young, ambitious TCNs, the marginal attractiveness of each additional transfer dropped considerably, for both personal and family reasons. With advancing age, the desire to settle down increased; this was coupled with the need for a stable primary group for the TCNs and for their growing children. For the PCNs, in comparison, the return to HQ was instrumental in solving these problems, or at least making them less acute.

Transfer. For PCN executives in MNCs of the dominant type, a transfer from one country to another gave rise to four sources of anxieties: time, place, position, and independence. The lack of manpower planning made it impossible for them to know when they would be transferred, whether the transfer would be HQ-bound or subsidiary-bound, and what position was awaiting them in their new assignment. This situation caused anxiety for the PCNs, because they could not plan ahead. In addition, transfers from subsidiaries back to HQ usually involved a loss of managerial independence, especially for those PCNs holding top level positions abroad. Even when their reentry entailed a promotion, they would have to readjust to working as part of a closely knit team at HQ, with more intensive and immediate scrutiny of their daily activities.

TCNs shared with PCNs the first three sources of transfer anxieties. The question of independence was not salient, because for them transfer was always to another subsidiary. On the other hand, TCNs reported a unique source of anxiety not shared by the PCNs: the lack of a home base to return to. A small number could serve certain periods at HQ, if they wanted to do so. However, since top positions were not open at HQ, TCNs were reluctant to leave top and relatively independent positions in the subsidiaries for upper middle level positions at HQ in which they would be less independent. In addition, HQ was considered as one more transitory assignment away from their own country. Consequently, TCNs considered it in their best interest to establish themselves

[2] Throughout the article, presentation of scores follows the same order: HCNs first, followed by TCNs.

at the head of a subsidiary in one particular third country, and to perpetuate themselves in it as long as possible, preferably until retirement.

The HCNs in the subsidiary headed by the TCNs were dissatisfied with this tendency of their TCN superiors to stay permanently in the top jobs, since this blocked their own promotion. This was especially the case with those HCNs who found it difficult to accept transfers to third countries as a condition for promotion to top positions, or those not considered eligible by HQ for overseas assignments.

Income. In the MNCs of the dominant type, there were gaps in the income levels between the PCNs and the HCNs. Income was composed of two components: basic salary and fringe benefits. Basic salary was equal for both groups at each particular level. Fringe benefits, however, were allocated according to nationality. PCNs received several fringe benefits not common in the host countries; for example, free acquisition of the unique product or service of the MNC, car, participation in rent and education of the children, expense account, various types of insurance, and home leaves. These fringe benefits created a distinctive income gap between the PCNs and HCNs, and were a source of dissatisfaction for the HCNs, who were much less satisfied than the PCNs.

In the subsidiary headed by TCNs, both pay and fringe benefits were a function of nationality. The pay and fringe benefits for the HCNs were determined according to the compensation levels prevailing in the host country. The pay and fringe benefits for the TCNs were determined by HQ through personal contracts. The pay scales for the TCNs were much higher than those for the HCNs. On the other hand, while HCNs benefited from the fringe benefits common in the host country, the MNC did not offer these benefits to the TCNs.

The TCNs were much more satisfied with their pay than were the HCNs (3·09 *vs.* 4·00). The TCNs admitted that their pay represented "to a great extent" their skills and efforts. However, the TCNs did not consider the high pay as adequate substitute for the hardships involved in their uprootedness and constant transfers. Hence, they did not feel that the MNC had "a real interest in the welfare and happiness of those who work in the corporation" (2·57) and were not convinced that "the company tried hard enough to improve their working conditions" (2·85).

Familiarity and adaptability. PCNs arriving in host-country organizations to assume top positions ordinarily required a period of adjustment to their new environment. Their unfamiliarity with the organization and its environment naturally led to managerial mistakes, such as the imposition of an offensive HQ managerial style, aloofness, non-participative decision-making, inadequate patterns of interaction with major environmental forces, and censoring upward and downward communication. The PCNs found it necessary to cover up their mistakes in order to justify their transfer to top positions, especially in those cases in which the senior HCNs were highly competent and well qualified for

Vol. 5
1977

these same positions; otherwise, they lost face. Such behavior was resented by HCNs, and left a residue of bitterness and doubts about PCNs' loyalty to the MNC.

The behavior of the TCNs was no exception. Contrary to the expectations of the advocates of the genuine multinational model, the TCNs' experience in several countries prior to their arrival in the subsidiary had not turned them into effective multinational personalities. The TCNs reported the same kinds of adaptation difficulties previously reported by the PCNs. However, the intensity of their anxieties was much higher than that of the PCNs. The interviews revealed that their dysfunctional behavior stemmed from their unique status in the MNC. This status was viewed from two vantage points: their relationship with HQ, and their relationship with the HCNs. PCNs were much more certain of HQ backing and understanding, even when they made mistakes; TCNs, on the other hand, felt constantly tested under fire and did not expect to get the backing of HQ when they made mistakes. The PCNs felt that the HCNs resented the fact that they blocked their promotion. At the same time, the HCNs also recognized that the PCNs were functional as effective linkages with HQ; this was not the case with the feelings of the HCNs toward the TCNs. Therefore, the TCNs were under stronger pressure to justify their superior position than were the PCNs. As a result, their more intense anxieties increased their dysfunctional behavior.

Type of projects. PCNs on overseas assignments tended to concentrate on short range, non-risky and demonstration-type projects. Their criteria for choosing projects were determined more by private considerations than by organizational needs. They were eager to demonstrate to HQ their ability to adapt quickly to foreign environments, to diagnose the subsidiary's major problems and grasp its opportunities, and to produce impressive, highly visible and measurable results. This attitude was motivated by HQ's system of assessing its expatriates by such criteria as organizational growth, net income, market share, and operation within budgetary limits—which are highly visible and measurable—rather than by their ability to handle human problems within the subsidiary and those arising in the subsidiary's interaction with its environment. The handling of such problems was an intricate, long-range effort, and the results were difficult to assess. Consequently, PCNs tended to neglect long-term projects requiring more thorough familiarity with local conditions and opportunities, or long range projects of organizational renewal which could increase the fit between the corporation's goals and the HCNs' needs. The HCNs considered such behavior by the PCNs to be both dysfunctional to the MNC and detrimental to their personal interests.

Our extensive interviewing of TCNs and their HCN subordinates and colleagues in various organizations revealed that TCNs, like PCNs in the MNCs of the dominant type, tended to concentrate on short-range projects. Moreover, they felt under pressure to choose either projects conforming with

precedents approved by HQ, or projects whose results could be easily and quickly manifested. They were motivated by the same considerations motivating the PCNs, but their need to concentrate on short-range projects was more intense because of their unique status in the MNCs.

The study revealed that the PCNs developed a primary group relationship with their peers at HQ prior to their assignment abroad. During their assignment abroad, they were likewise able to develop primary group relationships with compatriots in the subsidiaries. After their return to HQ, they either renewed their previous relationships or developed new ones. Hence, for the PCNs, the cycle of transfer and reentry was accompanied by powerful primary group support. This fact diminished to some extent their dependence on HQ officials in charge of their career ladders. Our research points out that TCNs, however, because of their uprootedness and the transient character of their career, were unable to develop primary group relationships influential at HQ. This was so because the only primary group relationships they could develop were with HCNs and TCNs. Primary groups composed of HCNs or TCNs (or both) did not have the same power—and, therefore, could not offer the same kind of support—as did primary groups composed of PCNs. Lacking a primary group influential at HQ, the TCNs developed a much stronger sense of dependence on HQ officials than did the PCNs. While HQ was both a supportive primary group and a reference group for the PCNs, HQ was only a reference group for the TCNs. Under these circumstances, the TCNs had no choice but to follow what they considered to be HQ preferences. Therefore, they concentrated on short-range projects which could demonstrate their contribution to the MNC, and hence their indispensability.

Vol. 5
1977

Leadership style. The dependence of PCNs on HQ precipitated another source of morale problems common to MNCs of the dominant type: PCNs' tendency to import into the subsidiaries the managerial style prevalent at HQ. That style was not necessarily suitable to the needs of the HCNs, or congruent with their perceptions of what an effective managerial style should be.

In the subsidiaries headed by PCNs, HCNs were dissatisfied with the leadership style of their superiors. In particular, they were dissatisfied with the following leadership patterns: goal definition and goal evaluation, sensitivity to subordinates problems, friendliness and accessibility, personal example, encouragement of team formation and of outstanding work records, and coaching.

Most instructive was the fact that PCNs[3] adopted this leadership style even though they were convinced that less detached, more considerate leadership behavior was more appropriate [6]. Some adopted it because of their uneasiness about holding positions superior to HCNs who were no less (or even more) competent. Others felt superior because of their conviction, based on MNC

[3] In all of the subsidiaries headed by PCNs, the differences between PCNs' perceptions of their current leadership patterns and the desired ones were statistically significant.

policies, that their subordinate HCNs would never be allowed to reach top positions at HQ. Still others adopted the style out of a sense of superiority toward the HCNs nurtured by past international experience and accomplishments. Some adopted the style as a result of a combination of all three considerations.

One problem merits special attention: definition of goals and evaluation of goal achievements. The source of the problem was the fact that top officials at HQ and PCNs heading the subsidiaries tended to define organizational goals in broad strokes rather than in clear-cut terms, thus leaving room for flexibility in interpretation. This flexibility was a double-edged sword. It might have been welcomed as a good practice if it had not allowed for post-facto interpretation of the goals during goal evaluation. HCNs found it difficult to grasp the intent of the generally defined goals, and felt that their PCN superiors interpreted the goals post facto in a manner slighting HCNs' achievements.

Similarly, the TCNs heading the subsidiary tended to impose what they considered to be the style desired by HQ. The TCNs believed that their management style was appropriate under the circumstances, and the responses of the HCNs to the questionnaire indicated that this was indeed the case. Following are the scores of the HCNs on their attitude concerning their superiors' leadership patterns: goal definition and goal evaluation practices (3·62); sensitivity to subordinates' problems (3·69); friendliness and accessibility (3·91); personal example (4·02); encouragement of team formation (3·63) and of outstanding work records (3·63); and coaching (3·37). These scores, which were validated in the personal interviews, were higher than the scores of the HCNs in the subsidiaries headed by PCNs.

The level of satisfaction of the TCNs was similar to that of the HCNs (3·66 *vs.* 3·44, mean of means of 22 questions dealing with different aspects of leadership).

Two morale problems were particularly disturbing to the TCNs. Like the PCNs, they were especially dissatisfied with the fact that goals were defined by HQ officials in general terms and were disadvantageously interpreted post-facto (3·14). Their satisfaction was significantly lower than that of their PCN counterparts in the other subsidiaries. The TCNs' satisfaction was also lower than their HCN subordinates, because the TCNs' practice of goal definition and evaluation was more compatible with the expectations of their HCN subordinates. Second, the TCNs were also highly critical of insufficient efforts by HQ to encourage team formation and teamwork within their ranks, and were convinced that HQ encouraged competition among themselves (3·57).

Decision making and communication. In the MNCs of the dominant type, the detached leadership behavior of the PCNs heading the subsidiaries limited the participation of HCNs in decision making, and obstructed the free flow of communication between the two groups. These practices were expressed in the following patterns: making decisions at levels where the most adequate and

accurate information was not available; making decisions without consulting those affected; failure to use group meetings for solving work-related problems; inadequate interdepartmental exchange of information; inadequate intra-work group exchange of information; and disruption of upward and downward communication. Both PCNs and HCNs were dissatisfied with these patterns. However, the PCNs were much less so, and they did not initiate any planned change activities to change these patterns [4].

The findings in the subsidiary headed by TCNs were instructive in three respects. First, the level of satisfaction of the TCNs was lower than that of their counterparts in the other MNCs who were PCNs. Second, the levels of satisfaction of the HCNs and the TCNs in the subsidiary were similar (see Table 1, below), with two exceptions. Third, in these two cases in which there was a

TABLE 1. DECISION MAKING AND COMMUNICATION

	HCNs	TCNs
Decision making levels congruent with availability of adequate and accurate information	3·16	3·14
Prior consultation with those affected by decisions	2·27	2·71
Sharing of information	2·76	3·00
Use of group discussions for group decisions	2·95	3·14
Intra-group exchange of opinions and ideas	3·37	3·57
Interdepartmental flow of communication	3·09	3·00
Providing group members with essential work-related information	3·88	3·28
Flow of information for improving job performance	3·69	3·71
Upward transfer of information	4·07*	2·86
Downward transfer of information for effective group decision	3·78*	3·00

* Differences marked with asterisks were significant at the 0·05 level.

statistically significant difference, the direction of the difference was the opposite of the one found in the subsidiaries headed by PCNs: whereas the PCNs were more satisfied than their HCN subordinates, the TCNs who headed the subsidiary were *less* satisfied than their HCNs subordinates—a unique phenomenon.

The two items which the TCNs were significantly less satisfied were upward communication (4·07 *vs.* 2·86) and downward communication (3·78 *vs.* 3·00). In the interviews, it was revealed that these findings reflected TCNs' dissatisfaction with both HQ and their HCN subordinates. The TCNs were convinced that HQ refrained from providing them with confidential information and from having them participate in critical decisions, especially those concerning politically controversial issues. The TCNs were also well aware of the fact that the HCNs were deliberately withholding information essential for making effective decisions. In the interviews, the HCNs did not deny the charge; they explained that their behavior was motivated by their grievances concerning the TCNs' dysfunctional managerial patterns, and their conviction that the TCNs could not perform the role of effective linking pins between the subsidiary and HQ.

Vol. 5
1977

CONCLUSION

In the introduction, we presented five related assumptions of the advocates of the genuine multinational model concerning the TCNs. Our data, focusing on morale problems, enable us to shed light on three of those assumptions (numbers 3, 4 and 5). Our findings reveal that TCNs with high potential and desire for promotion are not always anxious to serve abroad, especially after one or two transfers; those not anxious to be transferred try to establish themselves at the head of a subsidiary in a third country and perpetuate themselves in it as long as possible, preferably until retirement. In addition, with regard to assumptions 4 and 5, the accumulation of variegated international experience is not necessarily translated into managerial behavior compatible with environmental needs and HQ's expectations. The data point out that TCNs, quite often unwillingly, tend to prefer HQ's expectations and disregard the subsidiary's needs.

It may be argued that the organizational dysfunctions and morale problems found in the subsidiary headed by TCNs are not inherent in the genuine multinational model, since the MNC we studied only *approximates* the ideal model. Our subject deviates from the genuine multinational model in one respect: its undeclared policy that TCNs cannot assume top positions at HQ. This argument is persuasive so long as no MNC currently in operation is identical with the ideal model. However, no far-reaching changes in the level of morale may be expected if MNCs change this undeclared policy. This is so because opening the road to the very top of the MNC to the most qualified employees, irrespective of nationality, could eliminate only one problem from among a cluster of problems related to blocked promotion. Because the number of positions at the very top is extremely limited, this opportunity will be a myth for most employees, although such a myth may be functional. The other six sources of morale problems, however, will still be present.

Our general conclusion is that each prevalent staffing policy has several sources of morale problems—some inherent in its uniqueness, and others shared by the alternative staffing policies—irrespective of the personal qualifications of the managers. MNCs should be well aware of the fact that no staffing policy is free of serious dysfunctional aspects. Hence, rather than contemplating possible staffing policies that would eliminate these dysfunctions, means should be sought for diminishing their effect. It is unrealistic to expect that a change toward the genuine multinational model will solve the basic morale problems prevalent in MNCs of the dominant type.

REFERENCES

1. COGAN RE (1973) Management development for nationals overseas. *Personnel* **50**(2), 24–30.
2. DYMSZA WA (1972) *Multinational Business Strategy*, Chapter 1. McGraw-Hill, New York.

3. HARARI E & ZEIRA Y (1974) Morale problems in non-American multi-national corporations in the United States. *Mgmt Int. Rev.* **14**(6), 45–53.
4. HARARI E & ZEIRA Y (1976) Limitations and prospects of planned change in multinational corporations. *Hum. Relat.* **29**, (7) 659–676.
5. HEENAN DA (1970) The corporate expatriate: assignment to ambiguity. *Columbia J. Wld Bus.* **5**(3), 49–54.
6. HOBBING E (1971) The World corporation: a catalytic agent? *Columbia J. Wld Bus.* **6**(4), 45–51.
7. HUNTINGTON S (1973) Transnational organizations in World politics. *Wld Politics* **25**(3), 333–368.
8. INTERNATIONAL CHAMBER OF COMMERCE, Ed. (1974) *Realities: Multinational Enterprises Respond on Basic Issues.* International Chamber of Commerce, Paris.
9. KOSHIRO K (1974) Labor problems in Japanese enterprises abroad: Singapore and Malaysia. *Jap. Lab. Inst. J.* **16**(12), 18–27.
10. KUIN P (1972) The magic of multinational management. *Harv. Bus. Rev.* **50**(6), 89–97.
11. MULLER RE & BAMET RS (1975) *Global Reach: The Power of the Multinational Corporations.* Simon & Schuster, New York.
12. PERLMUTTER HV (1969) The tortuous evolution of the multinational company. *Columbia J. Wld Bus.* **4**(1), 9–18.
13. PERLMUTTER HV & HEENAN DA (1974) How multinational should your top managers be? *Harv. Bus. Rev.* **52**(6).
14. RICHMAN BM & COPEN M (1972) *International Management and Economic Development.* McGraw-Hill, New York.
15. ROBERTS BC (1972) Factors influencing the organization and style of management and their effect on the pattern of industrial relations in multinational corporations. In *Transnational Industrial Relations.* (Ed. GUNTER HAAS), pp. 109–132. Macmillan, London.
16. ROBINSON RD (1973) *International Business Management: A Guide to Decision Making*, Chapter 4. Holt, Rinehart & Winston, New York.
17. ROBOCK SA & SIMMONDS K (1973) *International Business and Multinational Enterprises.* Richard D Irwin, Homewood, IL.
18. *Trends* (1975) *Profile of a Chief International Officer.* Heidrick & Struggles.
19. WATSON CE (1972) The Brazilianization of US subsidiaries. *Personnel* **49**(2), 53–60.
20. WEINSHALL TD (1975) Multinational corporations—their study and measurement. *Mgmt Int. Rev.* **15**(4–5), 67–76.
21. VAID KN (1975) Resolution practices of multinational corporations in India. Japan Institute of Labor, 1975 Asian Regional Conference on Industrial Relations, Tokyo, Japan.
22. YOUSSEF S (1973) The integration of local nationals into the managerial hierarchy of American overseas subsidiaries: an exploratory study. *Acad. Mgmt J.* **16**(1), 24–34.
23. ZEIRA Y (1975) Overlooked personnel problems of multinational corporations. *Columbia J. Wld Bus.* **10**(2), 96–103.
24. ZEIRA Y (1976) Management development in MNCs. *California Mgmt Rev.* **18**(4), 34–42.
25. ZEIRA Y & HARARI E (1975) Planned organizational change in multinational corporations. *Advd Mgmt J.* **40**(3), 31–39.
26. ZEIRA Y, HARARI E & IZRAELI NUNDI D (1975) Some structural and cultural factors in ethnocentric multinational corporations and employee morale. *J. Mgmt Studies* **12**(1), 66–82.
27. ZENOFF DB (1969) *International Business Management*, Chapter 9. Macmillan, New York.

ADDRESS FOR CORRESPONDENCE: *Dr Y Zeira, Faculty of Management, The Leon Recanati Graduate School of Business Administration, Tel Aviv University, Ramat Aviv, Tel-Aviv, Israel.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977. Pergamon Press. Printed in Great Britain

New Directions for Management Accounting

EA LOWE

AM TINKER

University of Sheffield

(*Received May 1976; in revised form October 1976*)

Using a general systems rationale, this paper develops a theoretical structure for approaching the problem of management accounting. The management control problem is explicated in terms of maintaining a relationship between the enterprise's structure and its environment. An enterprise's structure is composed of three elements (and their inter-relations): a decision and control subsystem; a financial funds subsystem and an operating (physical transforms) subsystem. Portrayed in these terms, the problem of management accounting is shown to require a methodology which is able to take cognizance of economic, sociological, psychological and other aspects of the enterprise system. The model described here provides a general intellectual frame of reference for ordering the problem.

1. 5
977

1. INTRODUCTION

THE REAL tragedy of the Titanic was that the whole episode was avoidable! The unsinkable embodiment of engineering excellence sank for lack of a careful analysis of the ship's prospective environmental conditions in relation to her capabilities for remaining afloat.

Has the accountancy profession anything to learn from the loss of the Titanic in relation to the design of their education system? The career period of a newly qualified accountant is approximately 40 years. Although post-qualifying education and experience sometimes play an important role in an accountant's development, it is the expected demand conditions for 'accounting services' expected to prevail over this 40 year time horizon rather than current or past circumstances that define the accountancy education problem. Yet do we ever see an assessment of such expected demand conditions used to derive proposals for professional accountancy training? We think not. The rapid expansion of consultancy divisions of audit firms; the increasing importance of the industrial and commercial employment sector; greater competition from university graduates in economics, behavioural science, operations research, etc.; all such

trends in the accountancy profession's environment emphasise this change in longer-term demand for accounting skills. Fortunately some recent events suggest that the UK profession acknowledges educational change as a necessary first step in inducing revisions in accountancy 'technology' and practice. In particular, there is now evidence that the professional accounting bodies are making painstaking and deliberate progress towards proposals for changing accountancy practice [1, 11].

This paper is intended to contribute to the debate by describing a conceptual scheme of the management process that would not only enable the student to relate his accounting expertise to other views of management, but also provide a frame of reference for relating to more detailed proposals for accountancy training. This formulation is developed from a general systems perspective of the management process.

Thus our proposals here are directly relevant to the problems of training management accountants (or 'accountants in management'). Only partly do they apply to the problems of training accountants for professional practice, since the latter also requires a training which fully recognises *the socio-economic responsibilities of accountancy practice.*

Vol. 5
1977

2. A GENERAL SYSTEMS ANALYSIS

We can begin by recognising that management scientists (including management accountants), like any other (social) scientists, are not concerned with objects or entities in their entirety. Rather they abstract a particular viewpoint which depends upon the kinds of problems they are most immediately concerned about. A general systems viewpoint is concerned with identifying the whole of the system to which the problem relates and thus seeks to avoid taking that system apart in a manner which results in the loss of essential characteristics of the problem. The problem in this instance may be said to be the survival of a particular enterprise at a specified level of 'welfare'. This requires an understanding of those fundamental traits which contribute to its viability and adaptivity as a financial entity. In this respect accountancy practice and textbooks are grievously at fault, emphasising, as they do, the mechanistic and routine aspects of management processes at the expense of the fundamental innovative and imaginative aspects on which longer term financial viability and adaptivity depend.

The firm's behaviour may be described in terms of, for example:

An accounting viewpoint, e.g. different accounting profits associated with different budgets for the firm.

A moral or social viewpoint, e.g. number of employees satisfied with career

prospects; improvements in the firm's avoidance of polluting the environment.

A Government viewpoint: numbers of employees; taxes paid; accident rates; contribution to exports, etc.

An economic viewpoint: product prices; factors influencing cost levels; output level decisions, etc.

Many management theorists have argued that certain 'ideal-type' organisational structures exist that would serve all firms in all circumstances [7]. These 'ideal-types' often approximate to Weber's bureaucratic form of organisation. According to this thesis the entire problem of management resides in the organisation's arrangement within its 'own boundary'.

We argue precisely to the contrary: the appropriate form of organisation for realizing a firm's objectives is contingent upon its environmental conditions. From both a cybernetic and information systems viewpoint a formal matching as between the firm's 'substantial' environment and its internal organisation can be shown to be a requirement for a 'successful' organisation. (Remember the Titanic!)

The impact of environmental conditions, (represented by vector *A* in Fig. 1) the behaviour of the firm (represented by vector *B*), has been recognised in part by management accountants in, for instance, the distinction between controllable factors and uncontrollable factors as determinants of organisational performance. But to take such a specific instance for illustrative purposes this cost classification for control purposes can be said to have led accountants sadly astray since it results in a failure to understand the decision consequences of such a classification in the light of general systems theory. The pertinent message is that 'uncontrollable' factors are often critical because of their influence on

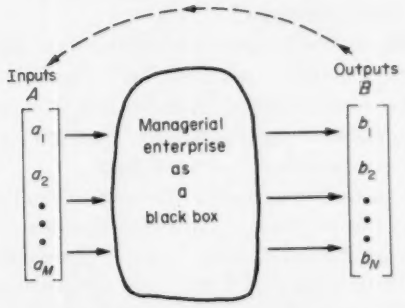


FIG. 1. *Matching the environment and organisation of an enterprise.*

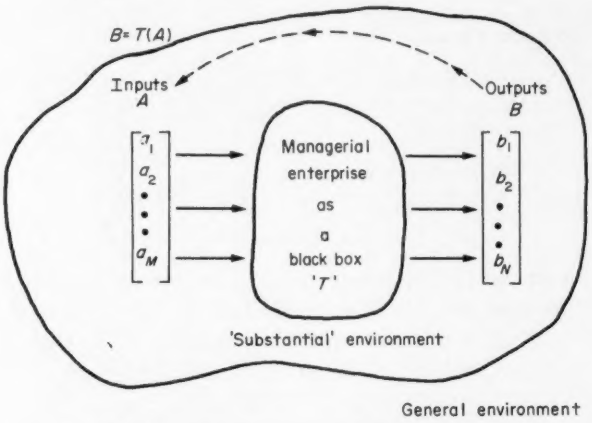


FIG. 2. *'Closure' in a system's analysis of an enterprise.*

the firm's well-being. Consequently the message implies that the firm's management should not assume that because some factor is uncontrollable, no action is required on their part. Clearly if the firm's objectives imply its survival, the uncontrollable factor requires to be 'controlled' by appropriate adaptation of the organisation. In many cases the 'control' of uncontrollable factors (sic) requires that other information relating to that factor is necessary for the firm's related decision-making processes. In practice, one often finds that the required information demands an improved ability to predict the changes in that environmental factor; often an improvement in the firm's forecasting faculty is called for.

Thus, at this general level of analysis, the behaviour under consideration (vector B), is a product of environmental conditions (vector A) and the organisation of the firm. That is: $B = T(A)$ where T is the transformation or organisation of the firm that relates environmental inputs to the behaviour of the whole organisation, in terms of its outputs.

At this level of analysis, the firm is regarded as a black box: the precise form of the transformation remains unspecified. Such a specification, in terms of a firm's technology, the deployment of resources through possible management policies, search rules, decision procedures, etc. requires further elaboration of the conceptual framework that is our concern here.

Given that aspects of the environment, relevant to explaining and predicting the behaviour of the organisation, have been identified, closure (in systems analysis terms) may be effected thereby, thus circumscribing the firm's 'substantial' (or relevant) environment.[1] These ideas are portrayed in Fig. 2.

[1] For discussions of the concept of closure in cybernetic modelling see Ashby [2] and Klir [6].

Vol. 5
1977

In Fig. 2 the closure represented by the boundary is intended to signify that all (perceived) organisational and environmental factors involved in explaining and predicting the behaviour of interest have been so included in the systems analysis. Prediction is thus a necessary condition for all management control processes although of course, it is not a sufficient condition.

In any particular systems analysis, the boundary is subject to a time horizon or a 'time resolution level' relating to the period of analysis. Clearly the longer the period of analysis, then the more likely the environment is to 'make its presence felt'. Consequently the more untenable may become many existing assumptions and the greater is the need to incorporate additional factors (and changes therein) explicitly into that analysis; such factors as the state of technology, political forecasts, changes in consumption patterns, etc. That is to say in Fig. 2 we draw the boundary correspondingly wider. Economic models have always attempted to recognise this aspect of management by a clear distinction, for instance, between short-term and long-term economic analysis. Unfortunately whilst the distinction is clear, the economic analysis of the short-term and long-run has not yet been made helpful to firms in an operational sense. Discounted cash flow methods probably represent our best attempt to do so, so far. But it is still quite evident that budgetary control methods, as presently generally practised and taught, have no formal means for including these notions in financial plans and budgets. This paper suggests that an operational distinction should be based on a specification of the whole causal mechanism intervening between various sets of inputs and outputs at a given time-accuracy resolution level. Figure 3 provides a general portrayal of this idea.

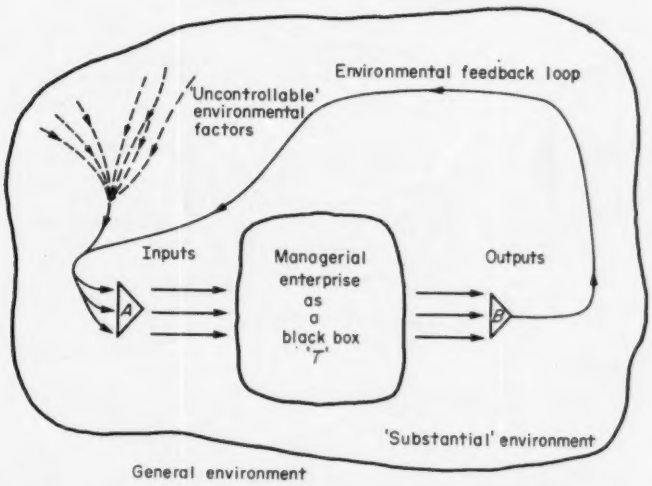


FIG. 3. *The 'casual' mechanism of enterprise stability.*

In an analysis of the behaviour of a firm through time, a feedback relation links behaviour (outputs) to environmental inputs. In these circumstances, the conventional accountancy distinction between controllable and uncontrollable influences becomes a most unhelpful one. For example, the price of capital to a firm may be determined by a market mechanism that is itself affected by information about and from that firm over a sufficiently longer time period. In such vital matters it is extremely difficult to determine the precise impact of a decision by the firm with current analytical financial techniques. But general systems theory, for instance, enables us to state clearly the relevant interdependencies and relationships.

3. A CONCEPTUAL DESIGN FOR ACCOUNTING FOR MANAGEMENT CONTROL

The general approach previously outlined can be formulated in terms of the conception of a 'managerial enterprise' shown in Fig. 4.

There it is suggested that the financial management problem may be investigated by focusing on the significant attributes of inputs, outputs, the internal structure of the system, and the relations between them. Thus inputs, defined as human needs, are represented as information, financial funds and physical resources. Similarly, information, financial funds and product flows are characteristic descriptions of the enterprise's output of satisfactions of human needs.[2]

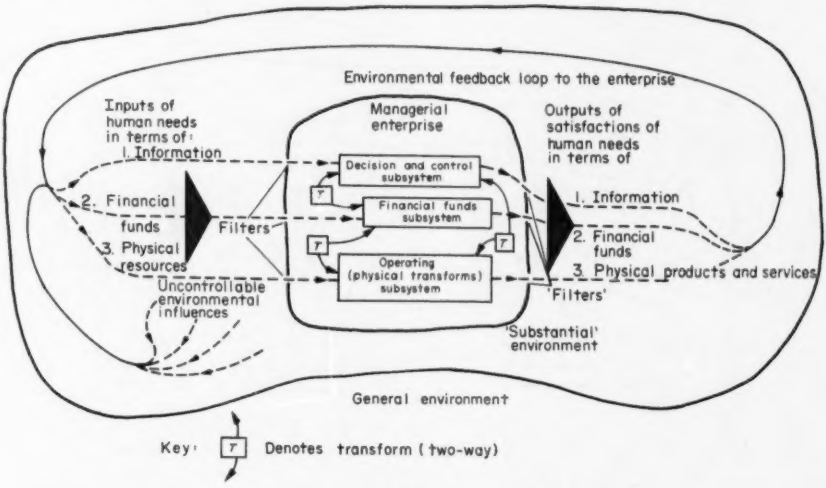


FIG. 4.

[2] This argument is developed by Lowe [8, 9], and Lowe and McInnes [10].

The internal structure of the enterprise in the illustration is portrayed as consisting of three interrelated principal elements, namely, the decision and control sub-system, the financial funds sub-system and the operating sub-system. These three elements are interrelated in that they are partial images of each other. In other words, input may be described by a vector with three components each of which describes the impact of that input upon each of the three sub-systems: the decision and control sub-system: the flow of funds sub-system: and the operating (physical transforms) sub-system. Conceptually the first is the most difficult to identify since it attempts to describe the 'informative' content of the input for the enterprise. It may be identified by its two principal component parts: a predictive model of the future possible (interdependent) states of the other two sub-systems; and, secondly, a model of choice that relates the feasible alternatives generated to the organisation's objective function.

Outputs in this model are the enterprise's responses to the stimuli (or inputs). They constitute the enterprise's attempts to influence the substantial environment and thus, in a cyclical sense, the next set of inputs to the enterprise. For instance, in our cost of capital example, it might be the capital market's revised view of the cost of capital to that firm, consequential upon the firm's changed policy with regard to matters such as the scope of its business or its means of raising capital or its dividend payments. Such changed policies might themselves be changed outputs (responses) in relation to previous inputs (stimuli) emanating from external, consumer and financial markets.

The decision and control sub-system, consisting of the 'brain' (the top managers) and the executive arm (those who execute formal policies, procedures and accounting control devices laid down by top management), is linked to the other two sub-systems. It is linked, firstly, to the financial funds element, thereby resulting in changes in the structure of funds according to criteria intended to achieve specific objectives; secondly, to the operating sub-system through a further two-way flow information-for-control linkage, thereby causing changes in production and production technology. In addition, the decision and control system contains management's 'imaginative faculty', for speculation and anticipation, involving both search for new opportunities and the reduction of hypothetical states to the consequences for the three sub-systems.

The financial aspects of the model, as a characteristic (or attribute) of the whole system, may be emphasized through the financial funds stock sub-system. Thus the behaviour of this sub-system might, for instance, feature a present value measure of plans in the form of a discount rate reflecting an opportunity cost of funds to organisational participants [8]. Recognition of this relationship, between the attribute and the whole system has of course important implications for management research and thus for company policy (see [12] and [4]).

Alternatively a behavioural science emphasis would involve focusing on the internal structure of the decision and control sub-system. This could be related to the financial funds sub-system through the monetary returns to conflicting

participants whose interests are represented by constraints and goals in the organisational decision system [4, 5].

The three sub-systems are circumscribed by an inner boundary which is permeable but monitored, and which contains the set of influence relations emanating from the co-operative character of the organisation, (for example the set of role relations, and authority and power relations). These directed social influences are thus reinforced by the tendency to resolve internal conflict by seeking a partial consensus; by resorting to pursuasion and legitimization to effect organisational change.

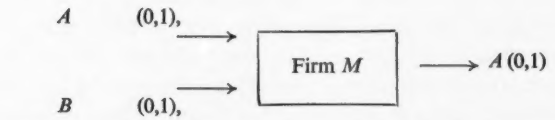
The use of this model can be briefly illustrated in relation to two fundamental kinds of management problems: problems of analysis and problems of synthesis [6]. Problems of analysis and synthesis may be defined with reference to this simple illustration.

4. AN ILLUSTRATION

A manager is interested in explaining the relationship between two input variables and one output variable of his firm *M*. Each of the input and output variables may take on only two states, (0,1) where

Input variable	State	Meaning of state
A	0,	High unemployment in the locality
	1	Low unemployment in the locality
B	0,	Competitors prices below those of firm *A*
	1	Competitors prices above those of firm *A*

Output variable	State	Meaning of state
W	0	loss
	1	profit



The systemic variety is $2^{2^2} = 16$. [3]

[3] For a discussion of systemic variety see [3].

The sixteen possible configurations of inputs and outputs (systemic or logical variety) are as follows:

A	B	W															
		1	2	3	4*	5	6	7	8	9	10	11	12	13	14	15	16
0	0	0	0	0	0	0	0	0	0	1	1	1	1	1	1	1	1
0	1	0	0	0	0	1	1	1	1	0	0	0	0	1	1	1	1
1	0	0	0	1	1	0	0	1	1	0	0	1	1	0	0	1	1
1	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1	0	1

From a large number of observations the manager notes from his empirical analysis that the firm behaves regularly according to transformation $W = 4$ above.

i.e.

A	B	W
0	0	0
0	1	0
1	0	1
1	1	1

To explain this regular pattern of relations between variables A, B and W, the manager needs to examine the internal structure of Firm A. The approach would be analogous to the procedure previously outlined: by defining different sets of elements of Firm A (e.g. departments, social groups, etc. representing the three sub-systems previously described) and examining the relations between characteristics of the elements and the inputs A and B and output W, the manager thus seeks a regular pattern that provides an explanation. For instance, after examining the internal structure of Firm A, the manager may discover that in times of high unemployment in the locality, output may be restricted by informal agreement to avoid competition for jobs. This explanation is consistent with the relations between A, B and W in transformation 4. Thus problems of analysis involve determining an explanation of the relations between inputs and outputs (e.g. transformation 4) with reference to a given structure.

In certain respects problems of synthesis are the converse of problems of analysis. Given a desired or prescribed behaviour (for example, $W = 1$ = profit in the illustration), the problem is to synthesize that organisational form that will realise the desired behaviour. Further analysis of the internal structure may reveal that certain arrangements of the decision and control sub-system (e.g.

employee payment schemes, redeployment programmes, etc.) affect the operating sub-system in a manner that realizes a satisfactory state of the financial funds sub-system for the firm in conditions of high unemployment. In this sense it may be possible to synthesise an internal structure, according to environmental conditions, that produces a profit.

This simple example perhaps serves to illustrate an approach to synthesizing an organisation in relation to a particular environment, and thereby changing output behaviour in accordance with the firm's objectives. It is important to note that, as in most management situations, the problems cannot be adequately resolved by a pre-occupation with the financial, personnel, marketing or production aspects alone. Such an approach would only do violence to a complex system of interrelated variables. Rather this total system must be respected, preserved as an entity and studied as a complete set of working relations through a general systems approach. The three sub-systems discussed previously provide a general method of classifying and relating these organizational variables.

CONCLUSIONS

Vol. 5
1977

The conceptual framework outlined in this paper may be regarded as an educational model for developing specific schemes for managing particular enterprises. It is not our purpose here to specify a way of programming the activity of a particular enterprise; rather we are concerned with the problem that is logically prior to this i.e. proposing a model to classify and order *ideas about ideas* for managing an enterprise in a coherent manner. This, we suggest, is the accounting educationalist's design problem. Thus it would be possible for an accounting student to enrich his view by considering, for instance, sociological, social psychological, quantitative, economic and other aspects of management problems, using such a model as the one presented here, as his 'intellectual framework for analysis'.

The meta-theoretic character of the model is particularly important from an educational viewpoint. The model is not a precise empirical theory, addressed directly to the workings of managerial problems; rather it provides a cognitive bridge for beginning to integrate ostensibly diverse theories and associated evidence and relating them to the problems of interest to us.

REFERENCES

1. Accounting Standards Steering Committee (1975) *The Corporate Report*. The Institute of Chartered Accountants in England and Wales.
2. Ashby R (1964) *Introduction to Cybernetics*. Methuen, London.
3. Beer S (1966) *Decision and Control: The Meaning of Operational Research and Managerial Cybernetics*. John Wiley, London.

4. BONINI CP (1967) *A Simulation of the Information and Decision Systems in the Firm.* Markham, London.
5. CYERT R & MARCH J (1963) *A Behavioural Theory of the Firm.* Prentice-Hall, Englewood Cliffs, N.J.
6. KLIR G (1969) *An Approach to General Systems Theory.* Van Nostrand, New York.
7. KOONTZ HD & O'DONNELL CJ (1964) *Principles of Management.* McGraw-Hill, New York.
8. LOWE EA (1970) The information content of financial statements, financial planning and management control: towards an integration. *J. Bus. Fin.* **2**(2), 3–11.
9. LOWE EA (1972) The finance director's role in the formulation and implementation of strategy: a framework for financial strategy. *J. Bus. Fin.* **4**(4), 58–70.
10. LOWE EA & MCINNES JM (1971) Control of socio-economic organisations: a rationale for the design of management control systems (Part 1). *J. Mgmt Studies* **8**(2), 213–227.
11. SOLOMONS D & BERRIDGE T (1974) *Long Range Enquiry into Education and Training for the Accountancy Profession.* Gee, London.
12. WILLIAMSON O (1964) *The Economics of Discretionary Behaviour: Managerial Objectives in the Theory of the Firm.* Prentice-Hall, Englewood Cliffs, NJ.

ADDRESS FOR CORRESPONDENCE: *Professor EA Lowe, University of Sheffield, Division of Economic Studies, Sheffield, S10 2TN*

. 5
77

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977. Pergamon Press. Printed in Great Britain

Optimal Design of a Post Office

KR BALACHANDRAN

University of Kentucky, Lexington, USA

(*Received May 1976; in revised form September 1976*)

A mathematical model which describes the postal operations is discussed in this paper. This paper is confined to the problems which arise within a post office in mail sorting and classifying operations.

1. 5
77

IN A post office, two objectives of fundamental importance in the processing and distribution of mail are decreasing costs and avoiding mail delays. It is sometimes possible to reduce both delay and costs by introducing new mail flow designs but, eventually, these goals can be achieved only at the expense of each other. Oliver and Samuel [1] have dealt with reducing delay subject to total cost. This paper deals with designing the post office to minimize total costs subject to restrictions on delay.

MAIL FLOW IN A TYPICAL POST OFFICE

The flow of two main forms of mail (letter mail and parcels) in a Post Office is described in this section. The less important forms of mail are not considered here beyond the initial culling stage where they are separated from the letter mail. This paper is written with a view toward future modeling of major postal operations, with a view to minimize total cost or mail delay; consequently, flow descriptions of less important forms of mail are omitted.

Parcel flow

Parcels are deposited into the post office by post office trucks and private company trucks who bring their parcels directly to the main post office. Incoming and outgoing parcels follow different routes. Air parcels and special handling parcels come in a separate package and are processed separately. The outgoing parcels, after they are delivered by truck, are stored on conveyors. Parcels are allowed to flow in small quantities for primary sorting, and conveyors

are again used to take them to the secondary sorting area. A small number of sample parcels from the company trucks are weighed for correct postage and then delivered to the secondary sorting. Local parcels are separated at primary sorting. Many of the transit parcels are handled in pouches.

In a mechanized system, the pouches are placed on moving trays and are directed to their destinations. The tray releases the pouch at the selected place, and the pouch joins an outgoing truck or railroad carrier. Loose parcels consist of outgoing parcels originating in or near the city, local parcels, and those parcels from pouch sorting which must be opened up for finer sorting. Two people work as a team on each tray conveyor for loose parcel sorting. The tray conveyor moves at a rate of 60 trays per minute, but it seems that, on the average, only every third tray is used; consequently, the number of parcels sorted is much lower. Local parcels are sent for secondary sorting. Parcels are packed in pouches and routed to trucks. Television receivers are used at various places to detect jams. Damaged parcels following any jam are repaired and then returned to the sorting area. Parcels of uneven shapes are processed manually. The entire operation is performed in three floors—the first floor for pouch sorting, the third floor for pouch shake-out, and the second floor for loose parcel sorting.

Vol. 5
1977

Letter mail flow

Outgoing mail operation starts with culling, where the airmail, local bulk, second class and other unusual mails are separated from ordinary first class mail. The ordinary first class mail is routed through an edger–stacker and is manually fed into a Mark II Facer–Canceller, where the stamps are recognized and cancelled, or through a modified Facer–Canceller, where the stamp phosphorescence levels are recognized and the stamps are cancelled. Airmail and local mail are separately hand stacked and run through a facer–canceller for cancellation. The local mail is sent by truck to another building, where the incoming mail is processed. Slugs, flats, and second class mail are cancelled on the conveyor itself manually and put in bins to be sent to sorting. Some of the ordinary first class mail is put in special green pouches to be sent by air, provided that air space is available. Ordinary first class letters are divided into 50 categories in the primary sorting and then sent for a secondary sorting. Airmail and special delivery mail is also processed through two sortings before being packed in pouches for out-of-town dispatches. Incoming mail flow is much simpler. Incorrectly collected mail is stored in a bin and sent back to join the outgoing mail. A third sorting is performed by the mailman himself to arrange the mail according to street address.

In a mechanized system the mail coming out of the Facer–Canceller is then hand delivered to a Letter Sorting Machine. This machine has 12 operators, each working at a rate of 60 letters per minute, with 240 bins for sorting. The letters are picked up manually from the bins and sent to secondary sorting, if

needed, or put in pouches. Secondary sorting is a small operation, since the Letter Sorting Machine has high capacity for sorting up to a city in many cases. The operators and the people who pick up letters from bins exchange positions periodically to avoid fatigue.

DEFINITIONS

Stages

A stage in a post office is defined as an operation whose design is independent of the designs of all other operations, for normal working. For example, the cancelling operation and the sorting operation can be considered in distinct stages—whereas, the primary and secondary sortings cannot be so considered, since the amount of automation in one sorting affects the required number of sorts in the other sorting. Another example is the mechanized edger–stacker, which is not used if letters are manually cancelled. The edger–stacker consequently is not a separate stage. Thus, letter processing consists of culling, cancelling, sorting, and storing stages.

Time periods

From the mail receipt distribution graph (Fig. 1), it is assumed that post offices work primarily from 6:00 a.m. to 12:00 p.m. in two 8-hr shifts. Each 8-hr shift is one time period. More consideration should be given to establish proper time periods for each post office. It is expected that the time periods will depend on mail receipt distribution, working hours of the post office, outgoing mail dispatch time (i.e. the timing of trains, trucks, and airplanes) and the dispatch capacities.

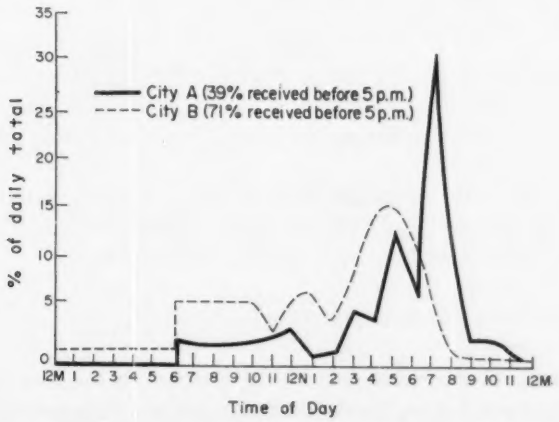

35
30
25
20
15
10
5
0
% of daily total
City A (39% received before 5 p.m.)
City B (71% received before 5 p.m.)
12M 1 2 3 4 5 6 7 8 9 10 11 12N 1 2 3 4 5 6 7 8 9 10 11 12M
Time of Day

FIG. 1. *Typical hourly receipt of first class mail at processing locations.*

One time period may also be something other than one shift. Usually workers are hired for 8 hr shifts. There is a possibility of hiring part-time workers for shorter periods. In this case, the time period should be redefined to suit the situation. The consequent increase in the size of the mathematical model will not be prohibitive.

PROCESSING RATES OF STAGES

The required processing rates will depend on the time periods, input distribution, and accepted delay of letters. The processing rates should be estimated so as to reduce the idleness probability of all stages. This is achieved by having the highest requirement for processing rate at the first stage and reducing the rates successively at the following stages. That the idleness probability is least with such a design is established in the theory of queues. The problem is to obtain the required processing rate at the first stage through using the mail receipt distribution and accepted delay of letters. One method is to find the area under the mail receipt rate curve and, within the time period, to find the height of a rectangle drawn with the duration of the time period as length and a certain proportion, Φ, of the calculated area as the area of the rectangle. The proportion Φ can be arbitrarily selected, depending on the dispatch timings and capacities. Areas in this discussion represent the amount of work performed, and it should be noted that the area of the rectangle does not necessarily represent the work performed during the corresponding time period. Letters cannot be processed before their arrival; consequently, the work performed during the time period considered is given by the intersection of the rectangle and the area below the mail receipt curve. Φ should be chosen so that the work performed is sufficiently suited to the dispatching process.

Vol. 5
1977

Zero-one integer programming formulation

For purposes of discussion let:

b_{jt} = required processing rate at stage j in time period t; $j = 1,2,\text{----},m$; $t = 1,2,\text{----},T$;
a_{ijt} = rate at which equipment i at stage j can work during time period t; $i = 1,2,\text{---}, n_j$;
c_{ijt} = fixed cost for using equipment i at stage j during period t;
f_{ij} = fixed cost per day for possessing equipment i at stage j;
x_{ijt} = 1 if equipment i is used at stage j during period t;
x_{ijt} = 0 otherwise;
y_{ij} = 1 if equipment i is used in stage j;
y_{ij} = 0 otherwise.

A superscript M on a variable refers to men rather than equipment.

It should be noted that the number of men employed during different periods may vary but equipment, once purchased, causes a cost of possession during all

periods whether or not it is used during a particular period.

Mathematically, the problem can be stated as follows:

$$\text{Minimize } Z = \Sigma[c_{ijt}\, x_{ijt} + c^{M}_{ijt}\, x^{M}_{ijt}] + \sum_{ij} f_{ij}\, y_{ij} \tag{1}$$

Subject to the following constraints for this problem:

$$\sum_{i}[a_{ijt}\, x_{ijt} + a^{M}_{ijt}\, x^{M}_{ijt}] \geq b_{jt},\ j = 1, ---, m;\ t = 1, ---, T. \tag{2}$$

$$\sum_{t} x_{ijt} \leq y_{ij} \tag{3}$$

$$x_{ijt},\ x^{M}_{ijt},\ y_{ij} = 0 \text{ or } 1.$$

Equation 1 gives the expression for total cost incurred per day by possessing and using equipment and employment of men. Equation 2 states the total processing rate achieved at each stage during each time period by using the particular number of men and equipment (as defined by the variables) to satisfy the requirements. Equation 3 implies that an item of equipment used in any period causes a state of possession during all periods.

In the preceding discussion, men, i, may consist of any number of men and the number need not be the same for all i. Wherever men must work together, they will be represented by a single subscript, i. This will reduce the size of the problem. An equipment item needs men to operate and maintain it. These men will not form part of the term x^{M}_{ijt}, since they do not directly work on the mail processing. The costs due to these men will be taken as part of operating and maintenance costs of the equipment.

COST DATA

The following cost data is needed to implement the preceding analysis technique:

(a) Equipment
   1. Depreciation costs on equipment.
   2. Repair and maintenance costs, including labor.
   3. Spare parts cost.
   4. Equipment operation costs, including wages and overhead burden (e.g. benefits of workers).

(b) Men (Those men manually processing mail, their wages, and the overhead burden).

All costs to be calculated per selected time period.

For a typical problem:
Number of periods $= T = 2$.
Number of stages $= m = 4 = j_{max}$.

Number of equipments at stage $j = n_j = 3 = i_{\max}$.
Number of men groups at stage $j = n_j^M = 3$.

Therefore, the total number of variables is 60. The total number of constraints is 20, besides the integer restrictions on the variables. For a discussion on integer programming, see Wagner [2].

ADDITIONAL APPLICATIONS

Parcel flow can be designed optimally, with the same model techniques. The other forms of mail (such as second class, oversized letters, etc.) do not seem to be undergoing any mechanical processing at the present time. However, when mechanical means are devised, the model described in this paper will facilitate obtaining an optimal design. The model can now be used to determine the number of men to be used at each stage for minimum total cost. The model is applicable for an entirely manual system also.

SEASONAL VARIATIONS IN DEMAND

The same model with the number of variables increased can be used to find the optimal design of a post office for processing during all seasons. The days during the Christmas season can be divided into two or more periods considering overtime, and the required processing rates at each stage can be calculated as before. In this case, there will be four or more time periods (if it is considered that the time periods may be different for normal and Christmas days). Consequently, the number of variables and the number of constraints will be increased when one is attempting to minimize yearly costs.

The size of the problem for letter processing will typically be:

Number of periods $= T = 4$.
Number of stages $= M = 4$.
Number of equipment items at stage $j = n_j = 5$ (approximate).

Therefore, the total number of variables is 180. The total number of constraints is 36, besides the integer restrictions on the variables.

ALLOWABLE COSTS FOR NEW EQUIPMENT

After finding the optimum design using currently available postal equipment, a further question as to the profitability of investing in a newly proposed equipment item can be answered. If the processing rate of the hypothesized

equipment is equal to the processing rate of an existing equipment item and/or group of men in the optimum design, the allowable cost is then simply the costs due to the equipment and/or group of men who can be replaced. The problem may not be so simple, however. Two difficulties may arise:

(a) The proposed equipment may do the work of two stages. In this case, the orientation of stages must be redefined.

(b) The expected processing rate of proposed equipment may not equal the processing rate of existing equipment and/or group of men. In this case, costs of both operating and possession of new equipment can be taken as unknown variables in the model described in earlier sections. If it is assumed that this new equipment must be used in any feasible solution, the problem can be solved to find the optimum design. The total cost for this design will be in terms of the unknown costs. Knowledge of the total costs for the existing optimum design and equating the two total costs will give allowable costs for the proposed equipment. Further information on the ratio between the operating and possession costs of proposed equipment will permit finding the allowable costs of operation and possession separately.

Thus, the preceding technique may be used to establish the economic feasibility cost for the development of new mail handling machinery with respect to optimum mail flow designs. The use of this technique, combined with equipment technical feasibility studies, should permit the selection of the most prolific mail handling techniques for early development.

REFERENCES

1. OLIVER RM and SAMUEL AH (1962) Reducing letter delays in post offices. *Ops Res.* 839–892.
2. WAGNER HM (1975) *Principles of Operations Research.* Prentice Hall, New Jersey.

ADDRESS FOR CORRESPONDENCE: *Professor KR Balachandran, University of Kentucky, College of Business and Economics, Department of Business Administration, Lexington, Kentucky* 40506, *USA*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977, Pergamon Press. Printed in Great Britain

Conservative Linear Programming with Mixed Multiple Objectives

AL SOYSTER
Virginia Polytechnic Institute and State University

B LEV
Temple University, Philadelphia, PA

DI TOOF
Arthur Young, Washington, DC

(Received July 1976; in revised form October 1976)

l. 5
77

In an ordinary linear program a single objective vector is constructed and one attempts to choose a decision vector to optimize this objective. Often multiple criteria exist or exact estimates for the components of a single objective vector are not entirely clear. For these cases a conservative decision-maker may want to choose an alternative that maximizes the objective value under the worst foreseeable circumstances. Herein we develop a unified framework for applying the maximin criterion to problems with various degrees of uncertainty attached to the objective vector. Three cases are solved via linear programming: (1) Complete Information, (2) Partial Information, and (3) Total Ignorance. It is shown that the functional value of the maximin solution decreases in a convex manner with increasing uncertainty. In addition certain relationships between maximin and efficient solutions are provided. Finally, an extension to integer constrained decision variables is presented.

ONE OF the usual formulations of an ordinary linear program is max $c \cdot x$

$$\text{s.t.} \quad Ax \leq b, A \text{ is mxn} \qquad \text{(LP)}$$
$$x \geq 0.$$

In this paper we shall consider the problem in which multiple objective criteria exist for (LP). Hence instead of a single objective vector c, there will exist a finite set of objective vectors C, where

$$C = \{c_1, c_2, c_3, \ldots, c_r \text{ and } c_i \in R^n\}.$$

In [7] Roy gives a comprehensive review of the many different approaches that researchers and practitioners have advanced for the optimization problem

with multiple criteria. One well known approach to the linear program with multiple criteria is simply to weight their objective functions according to weight $\{\lambda_i\}$ and use the composite objective function $\sum_{i=1}^{r} \lambda_i c_i$. The notion of efficient points for the multiple criteria linear program continues to attract considerable attention. A feasible solution $\bar{x}$ to $Ax \leq b$, $x \geq 0$, is said to be efficient (for maximization) if there does not exist another feasible solution x' such that $c_i \cdot x' \geq c_i \cdot \bar{x}$ for each i and strict inequality for at least one i. Construction and other notions of efficient points are considered in Charnes and Cooper [2], Geoffrion [4], Philip [6] and Evans and Steuer [3]. Recently certain strategies have been advanced that permit the decision-maker to progressively determine the relative utilities of the multiple objective functions. An example of this approach is given in [5].

A conservative approach that will be the subject of this paper belongs to the class of procedures described by Roy as "maximum reduction of uncertainty and incomparability." In this conservative approach we shall seek solutions to $Ax \leq b$, $x \geq 0$, that guarantee the largest possible return when various degrees of information are known about the elements of the objective set C.

There are at least two basic interpretations of the set C. First we can consider a firm in which there are r different objectives such as profit, growth, market share, employee stability and many others. Here an optimal solution for short run profit may fare badly with respect to growth and it is difficult to choose a vector x that is somehow optimal for the whole system. A second interpretation of the set C is that there is a single goal but for some reason the exact objective vector is not known. For example, the goal may be the maximization of profit but since some factors of the profit are not known with certainty, the exact components of the objective vector are also uncertain.

The second interpretation is the one more closely aligned with the standard framework of stochastic programming. Of course, some type of assumption or information concerning the relative likelihood of each of the objective vectors would be required. We shall describe herein a conservative strategy for the multiple objective problem that will apply to either of the foregoing situations.

$$\text{Define the set } C' \text{ as } C' = \{c \in R^n \mid c = \sum_{i=1}^{r} \lambda_i c_i, c_i \in C, \Sigma\lambda_i = 1, \lambda_i \geq 0, L_i \leq \lambda_i \leq U_i\}$$

where L_i and U_i are lower and upper limits on λ_i. We, of course, assume that λ_i exist that satisfy these conditions. A conservative solution for the multiple objective problem is a maximin solution to

$$\begin{array}{l} \max \{\min c \cdot x\} \\ Ax \leq b \; c \in C' \\ x \geq 0 \end{array} \quad . \qquad \text{(MM)}$$

Vol. 5
1977

The interpretation of (MM) is as follows: for each feasible $\bar{x}$ compute $c \cdot \bar{x}$ for each $c \epsilon C'$. There exists a $\bar{c} \epsilon C'$ which makes $\bar{c} \cdot \bar{x}$ a minimum, i.e. the worst possible c if $\bar{x}$ is chosen. The optimal decision vector x^* is the feasible x which makes the worst possible outcome the largest. If x^* is the optimal solution to (MM) with maximin value v^* this means that if x^* is used as the decision vector then a return of at least v^* is guaranteed.

The set C' is meant to include both of the earlier interpretations given to the multiple objective criteria problem. In the first case there actually exist r distinct goals that simultaneously bear upon the optimal choice of the decision vector. Here the intervals $[L_i,U_i]$ represent the uncertainty in the weight that should be accorded to goal i. For the second case a different interpretation is appropriate. Again r different criteria are assumed to exist but in this case only one of them will actually occur. The vector $(\lambda_1, \lambda_2, \ldots \lambda_r)$ is a random quantity where the domain of the density is bounded within the hyper-rectangle defined by $\lambda_i \epsilon [L_i,U_i]$. It follows that no matter the form of the density function that $L_i \leq E(\lambda_i) \leq U_i$, and furthermore, by an appropriate choice of density function one can have $E(\lambda_i)$ anywhere within $[L_i,U_i]$. Note that in this second case C' would represent the set of all possible expected values for the objective function so that (MM) is properly viewed as maximizing the worst possible expected value.

(1) PROBLEM IDENTIFICATION, FORMULATION AND SOLUTION

In the introduction a multiple objective linear programming problem is considered for which r different objective vectors exist although some information concerning the relative weighting of each criteria is known. While no information is given on the distribution of λ_i, it is known that the weighting factor λ_i for objective vector c_i lies somewhere in the closed interval $[L_i,U_i]$ as specified by C'. In general there are three levels of information that might be associated with the set of possible objective vectors defined by C':

(1) Total Ignorance;

(2) Partial Information;

(3) Complete Information.

For the case of Total Ignorance $[L_i,U_i] = [0,1]$ for each i so that absolutely nothing is known about the relative likelihood or subjective weight associated with each objective function. Thus the model of case 1 may be formulated as:

$$\begin{aligned} &\max \alpha \\ &c_i \cdot x \geq \alpha \qquad i = 1,2,\ldots r. \\ &Ax \leq b \\ &x \geq 0. \end{aligned}$$

Conversely, for the case of complete information each $L_i = U_i$ so each weight λ_i is a deterministic quantity. The formulation for case 3 is then

$$\max \sum_{i=1}^{r} \lambda_i c_i$$
$$Ax \leq b$$
$$x \geq 0.$$

Both these models are linear programs.

Now consider the case of Partial Information wherein some or all the parameters L_i and U_i are arbitrary in the unit interval. A linear programming formulation for case (2) is not so evident. In an early paper by Wolfe [10] such a formulation is obtained in the context of two person polyhedral games. A polyhedral game is specified when player one must choose a strategy, x, from $X = \{x | Ax \leq b,\ x \geq 0\}$, A is mxn, and player two chooses a strategy λ, from $Y = \{\lambda | \lambda F \geq g, \lambda \geq 0\}$ F is rxl. Let the payoff matrix, $\hat{C}$, be defined by $\hat{C} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_r \end{pmatrix}$ $\hat{C}$ is rxn · Wolfe considers the pair of problems:

$$\max_{x \in X} \min_{\lambda \in Y} \lambda \hat{C} x \qquad \text{(W-I)}$$

and

$$\min_{\lambda \in Y} \max_{x \in X} \lambda \hat{C} x. \qquad \text{(W-II)}$$

The main result in [10] is that if both (W-I) and (W-II) are feasible then their objective values are equal. Moreover, in this case the optimal solutions to (W-I) and (W-II) can be obtained by solving the linear programming problems

$$\max \quad g \cdot u \qquad \text{(LP-I)}$$
$$Ax \leq b$$
$$-\hat{C}x + Fu \leq 0,$$
$$x, u \geq 0$$
$$\min \quad y \cdot b \qquad \text{(LP-II)}$$
$$yA - \lambda \hat{C} \geq 0$$
$$\lambda F \geq g$$
$$y, \lambda \geq 0$$

This result shows precisely how a generalized version of (MM) under partial uncertainty can be solved via linear programming. Instead of C' consider arbitrary linear constraints on the variables, λ,

$$C'' = \{c \in R^n \mid c = \sum_{i=1}^{r} \lambda_i c_i,\ \lambda F \geq g, \lambda \geq 0\}.$$

Here the optimal solution to (MM) is obtained by solving (LP-I).

Vol. 5
1977

(2) SOME GLOBAL ANALYSIS OF UNCERTAINTY

In this section we investigate the impact of modifying the domain of the set of objective functions for (MM). In particular we shall be interested in the manner in which the objective function value varies as the set of objective functions for (MM) 'shrinks' or 'grows'. For example, consider the case of Complete Information wherein point estimates $\{\bar{\lambda}_i\}$ are given for each c_i. How does the objective value of (MM) change as these point estimates are enlarged to intervals $[L_i, U_i]$ and ultimately to [0,1]? The underlying motivation for such an analysis is to give the conservative decision-maker a tool for measuring the impact of uncertain information about the multiple objective functions.

One would naturally expect that as the domain of possible objective vectors increases that the objective value of (MM) would be non-increasing. Let C_1, C_2 be arbitrary closed convex sets in R^n such that $C_1 \subseteq C_2$ and define, for $\mu \epsilon [0,1]$,

$$f(\mu) = \max \min_{Ax \leq b} c \cdot x \quad c \epsilon (1-\mu) C_1 + \mu C_2. \tag{I$_\mu$}$$

The function $f(\mu)$ represents the value of the conservative maximin solution to (MM) when the possible objective vectors are contained in the convex vector sum $(1-\mu) C_1 + \mu C_2$. Note that when C_1 contains a single vector $\bar{c}$

$$\text{and } C_2 = \{c \epsilon R^n \mid c = \sum_{i=1}^{r} \lambda_i c_i, \ \sum_{i=1}^{r} \lambda_i = 1, \ \lambda_i \geq 0\}$$

then $f(\mu)$ represents the value of the conservative solution as one passes from Total Ignorance to Complete Information in a proportional fashion. A proof that $f(\mu)$ is convex and non-increasing is given in the Appendix.

The function $f(\mu)$ represents the value of maximin solution as the set of possible objective vectors varies between C_1 and C_2. A graph of this function would provide a decision-maker a means for evaluating marginal benefits of additional information concerning the set of possible objective functions. As the set of possible objective functions shrinks the maximin solution can only increase. But since $f(\mu)$ represents the value of an optimization problem, a complete graph of this function on the interval $\mu \epsilon [0,1]$ is not, in general, easily obtained. In lieu of the function $f(\mu)$ consider another function $g(\mu)$ that is everywhere a lower bound for $f(\mu)$, i.e. $g(\mu) \leq f(\mu)$ for all $\mu \epsilon [0,1]$. Define

$$g(\mu) = (1-\mu)^2 f(0) + \mu^2 f(1) + \mu(1-\mu) \left[\min_{c_1 \epsilon C_1} c_1 \cdot x_2 + \min_{c_2 \epsilon C_2} c_2 \cdot x_1 \right]$$

where x_1 is an optimal solution to (I_0) and x_2 an optimal solution to (I_1). Observe that the function $g(\mu)$ is a closed form quadratic defined on the unit interval. In the Appendix it is shown that $g(\mu)$ is a quadratic, convex lower bound to $f(\mu)$. The function $g(\mu)$ provides the decision-maker a global picture of a lower bound to the maximin solution $f(\mu)$. (Note that since $f(\mu)$ is convex that $(1-\mu) f(0) + \mu f(1)$ is an upper bound to $f(\mu)$.)

(3) RELATIONSHIP WITH EFFICIENCY

In the absence of additional or secondary information concerning multiple objective functions one must often be content with the construction of the set of efficient points. The concept of efficiency is of course useful in the cases of total ignorance and partial information, i.e. the exact role of each objective vector is not precisely known. A feasible solution to the multi-objective linear program, $\bar{x}$ is said to be efficient if there does not exist another feasible solution, x', with the property that $c_i \cdot x' \geq c_i \cdot \bar{x}$ for $i = 1,2 \ldots,r$, with at least one inequality holding as a strict inequality.

The set C'' in the definition of (MM) concerns r different linear criteria and the question arises whether a given maximin solution is necessarily efficient. The answer is no. This can be seen by considering case (iii) p. 60 of [3]. Let C'' be constructed from the convex hull of $c_1 = (1,0)$, $c_2 = (0,1)$, and take $A = (0,1)$ and $b = (1)$. The set of optimal solutions to (MM) would be $\{(x_1,1)|x_1 \geq 1\}$ but for this example there are obviously no efficient solutions. This example of course also demonstrates the existence of problems where no maximin solution is efficient. It is true, though, that if there exists a unique optimal solution to (MM) that this solution will be efficient. This result is proven in the Appendix. An immediate consequence of this fact is a possible means for choosing one particular member from the usually large set of efficient points. In general one would favor an efficient point that is also a maximin solution. Observe that Wolfe's results provides a formal strategy for determining the set of all maximin solutions which are efficient. All maximin solutions to (MM) are necessarily (part of) optimal solutions to (LP-I). Hence the set of all maximin solutions to (MM) is the solution set S where $S = \{x|Ax \leq b, -\hat{C}x + Fu \leq 0, g \cdot u \geq g \cdot u^*, x, u \geq 0\}$ where $g \cdot u^*$ is the optimal value of (LP-I). Methods for determining $E \subseteq S$ where E is the set of efficient solutions relative to a feasible set S, are discussed for example, in [3] and [11].

Consider now the inverse question: for what region of objective vectors is an efficient solution also a maximin solution? If x^* is efficient then there exists a vector $\lambda > 0$ such that x^* is optimal for

$$\begin{aligned} \max\ & \lambda \hat{C} x \\ & Ax \leq b \\ & x \geq 0. \end{aligned} \tag{e-I}$$

The dual linear program is

$$\begin{aligned} \min\ & y \cdot b \\ & yA \geq \lambda \hat{C} \\ & y \geq 0 \end{aligned} \tag{e-II}$$

Observe that via the fundamental objective function equality between (*e*-I) and (*e*-II) that x^* is optimal for (*e*-I) for $\lambda \epsilon \Lambda = \{\lambda|\exists y$ such that $yA \geq \lambda \hat{C}$ $y \geq 0$, $y \cdot b = \lambda \hat{C} x^*$, $\Sigma \lambda_i = 1, \lambda \geq 0\}$. The conditions in Λ guarantee a feasible

Vol. 5
1977

solution, y, to (e-II) for which $y \cdot b = \lambda \hat{C} x^*$, i.e. x^* is optimal for any $\lambda \epsilon \Lambda$ so that obviously x^* is a maximin solution for at least $\lambda \epsilon \Lambda$.

The question arises whether an efficient solution x^* can be a maximin solution relative to a larger set $\bar{\Lambda} \supset \Lambda$. The answer is yes. Consider the following example where the feasible region is $x_1 + x_2 \leq 1, x_i \geq 0$.

$$\hat{C} = \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} = \begin{pmatrix} (1,0) \\ (0,1) \\ (-1,-1) \end{pmatrix}$$

Observe that (0,0) is an efficient point and that (0,0) is optimal for

$$\begin{aligned} \max\ & (\lambda_1 c_1 + \lambda_2 c_2 + \lambda_3 c_3) \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ & x_1 + x_2 \leq 1 \\ & x_1, x_2 \geq 0, \end{aligned}$$

when $\lambda_3 \geq 1/2$. But (0,0) is the maximin solution for

$$\bar{\Lambda} = \{\lambda_1, \lambda_2, \lambda_3) \mid \Sigma \lambda_i = 1, \lambda_i \geq 0\}.$$

In the Appendix we show how to construct the largest set of objective vectors, say Λ^*, for which a given efficient point, x^*, will also be a maximin solution. This is illustrated via an example taken from [6]. In addition we construct the lower bound function to $f(\mu)$ for this example.

(4) AN INTEGER EXTENSION

In [1] Balas develops a duality theory for discrete programming. One convenient starting point for applying the results of Balas is the mixed integer program (P) as follows:

$$\begin{array}{lll} \max & & g \cdot u \\ Ax & & \leq b\,(y) \\ -\hat{C}x & + Fu & \leq 0\,(\lambda) \\ x \geq 0, & \text{integer} & \\ u \geq 0. & & \end{array} \qquad \text{(P)}$$

Let y be the n dual variables associated with the first n constraints and λ the r dual variables associated with the latter r constraints. For this particular problem the Balas dual problem (D) would reduce to the maximin problem:

$$\begin{array}{ll} \max & \{\min y \cdot b - (yA - \lambda \hat{C}) \cdot x\}. \\ x \geq 0, \text{integer} & \lambda F \geq g \\ & (y, \lambda) \geq 0 \end{array} \qquad \text{(D)}$$

A slight rearrangement provides:

$$\begin{array}{ll} \max & \{\min y\cdot(b - Ax) + \lambda\hat{C}x\} \\ x \geq 0, \text{integer} & \lambda F \geq g \\ & (y, \lambda) \geq 0 \end{array}$$

and since $y \geq 0$ it must be that $Ax \leq b$ or otherwise the inner minimization would be $-\infty$. Hence (D) is equivalent to:

$$\begin{array}{ll} \max & \{\min \ \lambda\hat{C}x\} \\ Ax \leq b & \lambda F \geq g \\ x \geq 0, \text{integer} & \lambda \geq 0 \end{array}$$

which can be conveniently rewritten with a change of variables as:

$$\begin{array}{ll} \max & \{\min c\cdot x\} \\ Ax \leq b & c\epsilon C'' \\ x \geq 0, \text{integer} & \end{array} \qquad \text{(D)}$$

Problem (D) is then equivalent to the maximin (MM).

According to Balas [1] Theorem 3, it follows that if either (P) or (D) is finite, then both are finite and their respective objective values are equal. Also included in Theorem 3 [1] are characterizations of possible duality states between (P) and (D) with respect to feasibility and unboundedness. But the principal result in [1] is summarized in the following:

Problem (P) has a finite optimal solution if and only if (D) has a finite optimal solution. Furthermore if either problem has a finite optimal solution then their respective optimal values are equal.

Another Theorem of Balas [1] implies that if (x^*, u^*) is an optimal solution to (P), then x^* is optimal for (D). Hence we can find an optimal solution to (MM) by solving (P).

Vol. 5
1977

ACKNOWLEDGEMENTS

We wish to thank the referees for their helpful comments on an earlier version of this paper.

REFERENCES

1. Balas E (1970) Duality in discrete programming. *Proceedings of the Princeton Symposium on Mathematical Programming* (Ed. Kuhn HW). Princeton University Press.
2. Charnes AW and Cooper WW (1961) *Management Models and Industrial Applications of Linear Programming*, Vol. 1. John Wiley, New York.
3. Evans JP and Steuer RE (1973) A revised simplex method for linear multiple objective functions. *Mathl Progr.* **5**(1), 54–72.
4. Geoffrion AM (1968) Proper efficiency and the theory of vector maximization. *J. Math. Analysis Applic.* **22**(3), 618–630.

5. GEOFFRION AM, DYER JS and FEINBERG A (1972) An interactive approach for multi-criterion optimization with an application to the operation of an academic department. *Mgmt Sci.* **19**(4), 357–368.
6. PHILIP J (1972) Algorithms for the vector maximization problem. *Mathl Progr.* **2**(2), 207–229.
7. ROY B (1971) Problems and methods with multiple objective functions. *Mathl Progr.* **1**(2), 239–266.
8. SOYSTER AL (1974) A duality theory for convex programming with set inclusive constraints. *Ops. Res.* **22**(4), 892–898.
9. Erratum to (8) (1974) **22**(6), 1279.
10. WOLFE P (1956) Determinateness of polyhedral games in *Linear Inequalities and Related Systems* (Eds. KUHN HW and TUCKER AW). Princeton University Press.
11. ZELENY M (1974) *Linear Multiobjective Programming*. Springer, New York.

ADDRESS FOR CORRESPONDENCE: *Professor B Lev, Department of Management, Temple University, School of Business Administration, Philadelphia, PA* 19122, *USA*

APPENDIX

Theorem 1

Let $C_1 \subset C_2$ be arbitrary closed convex sets in R^n. Then the function $f(\mu)$ is convex and non-increasing with respect to μ.

Proof

The fact that f is non-increasing follows directly from the inclusion $C_1 \subseteq C_2$. Let $\mu_1 \leq \mu_2$ be any two points in the unit interval and furthermore suppose that x_i is optimal for (I_{μ_i}). Choose $\lambda \epsilon [0,1]$ and note that

$$\begin{aligned} f[\lambda\mu_1 + (1-\lambda)\mu_2] = & \max_{Ax \leq b,\ x \geq 0} \ \min_{c \epsilon [1-(\lambda\mu_1+(1-\lambda)\mu_2)]C_1 + [\lambda\mu_1+(1-\lambda)\mu_2]C_2} c \cdot x \\ = & \max_{Ax \leq b,\ x \geq 0} \ \min_{c \epsilon \lambda\{(1-\mu_1)C_1 + \mu_1 C_2\} + (1-\lambda)\{(1-\mu_2)C_1 + \mu_2 C_2\}} c \cdot x. \end{aligned}$$

Now it is shown in Soyster [8, 9] that a dual problem for $f[\lambda\mu_1 + (1-\lambda)\mu_2]$ is

$$\begin{aligned} & \inf y \cdot b \\ & yA \subseteq \lambda\{(1-\mu_1)C_1 + \mu_1 C_2\} + (1-\lambda)\{(1-\mu_2)C_1 + \mu_2 C_2\} \qquad \text{(Dual)} \\ & y \geq 0. \end{aligned}$$

Since $f(\mu_1)$ and $f(\mu_2)$ are assumed finite then it follows that $f[\lambda\mu_1 + (1-\lambda)\mu_2]$ is finite for $\lambda \epsilon [0,1]$ so that $f[\lambda\mu_1 + (1-\lambda)\mu_2] = (\text{Dual})$. Let y_1, y_2 be dual, optimal solutions[1] corresponding to x_1 and x_2 for $f(\mu_1)$ and $f(\mu_2)$.

Observe that $\lambda y_1 + (1-\lambda)y_2$ is a feasible solution to (Dual) so that

$$f[\lambda\mu_1 + (1-\lambda)\mu_2] \leq [\lambda y_1 + (1-\lambda)y_2] \cdot b.$$

But since $f(\mu_i) = y_i \cdot b$ the convexity is proven.

[1] In the case optimal solutions for (Dual) do not exist one can always choose feasible y_i so that $y_i \cdot b$ is arbitrarily close to $f(\mu_i)$. A simple ϵ argument would complete the proof.

Corollary 1.1
If $x_1 = x_2 = x^*$ then x^* is optimal for $(I\mu)$ for all $\mu \epsilon \{\lambda\mu_1 + (1-\lambda)\mu_2 | 0 \leq \lambda \leq 1\}$.

Proof
Let $\nu^*(\lambda)$ be the maximin value generated by x^* for $\lambda \epsilon [0,1]$, i.e. define:

$$\begin{aligned}\nu^*(\lambda) &= \min c \cdot x^* \\ &\quad c \epsilon \lambda \{(1-\mu_1) C_1 + \mu_1 C_2\} + (1-\lambda)\{(1-\mu_2) C_1 + \mu_2 C_2\} \\ &= \min [\lambda c_1 + (1-\lambda) c_2] \cdot x^* \\ &\quad c_1 \epsilon \{(1-\mu_1) C_1 + \mu_2 C_2\} \\ &\quad c_2 \epsilon \{(1-\mu_2) C_1 + \mu_2 C_2\}\end{aligned}$$

Observe that now $\nu^*(\lambda)$ is equal to two separate terms

$$\begin{aligned}\nu^*(\lambda) &= \lambda \min c_1 \cdot x^* \qquad\qquad + (1-\lambda) \min c_2 \cdot x^* \\ &\quad c_1 \epsilon \{(1-\mu_1) C_1 + \mu_1 C_2\} \qquad\qquad c_2 \epsilon \{(1-\mu_2) C_1 + \mu_2 C_2\} \\ &= \lambda f(\mu_1) + (1-\lambda) f(\mu_2).\end{aligned}$$

Since f is convex it follows that

$$\nu^*(\lambda) \geq f[\lambda\mu_1 + (1-\lambda)\mu_2]$$

and the corollary is proven.

We next consider how a lower bound to the function $f(\mu)$ can be constructed. Since

$$f(\mu) = \max_{Ax \leq b,\ x \geq 0} \ \min_{c_1 \epsilon C_1,\ c_2 \epsilon C_2} [(1-\mu) c_1 + \mu c_2] \cdot x$$

it follows that

$$f(\mu) = \max_{Ax \leq b,\ x \geq 0} \{(1-\mu) \min_{c_1 \epsilon C_1} c_1 \cdot x + \mu \min_{c_2 \epsilon C_2} c_2 \cdot x\}$$

Now let x_1 be an optimal solution to (I_0) and x_2 an optimal solution to (I_1) and consider the feasible solution $(1-\mu) x_1 + \mu x_2$ for problem (I_μ).

It follows that

$$f(\mu) \geq (1-\mu) \min_{c_1 \epsilon C_1} c_1 \cdot [(1-\mu) x_1 + \mu \cdot x_2] + \mu \min_{c_2 \epsilon C_2} c_2 \cdot [(1-\mu) x_1 + \mu x_2],$$

and since the minimum of a sum is greater than or equal to the sum of the minima, for $i=1,2$

$$\min_{c_i \epsilon C_i} c_i \cdot [(1-\mu) x_1 + \mu x_2] \geq (1-\mu) \min_{c_i \epsilon C_i} c_i \cdot x_1 + \mu \min_{c_i \epsilon C_i} c_i \cdot x_2.$$

The RHS for $i = 1$ is

$$(1-\mu) f(0) + \mu \min_{c_1 \epsilon C_1} c_1 \cdot x_2$$

and for $i=2$

$$(1-\mu) \min_{c_2 \epsilon C_2} c_2 x_1 + \mu f(1).$$

Define

$$g(\mu) = (1-\mu)^2 f(0) + \mu^2 f(1) + \mu(1-\mu)[\min_{c_1 \epsilon C_1} c_1 \cdot x_2 + \min_{c_2 \epsilon C_2} c_2 \cdot x_1].$$

Vol. 5
1977

It follows that

$$f(\mu) \geq g(\mu) \qquad \text{for } \mu \epsilon [0,1].$$

Observe finally that g is also convex since the second derivative is non-negative,

$$g''(\mu) = 2\{f(0) - \min_{c_1 \epsilon C_1} c_1 \cdot x_2\} + 2\{f(1) - \min_{c_2 \epsilon C_2} c_2 \cdot x_1\}$$

and by definition of f both differences in the brackets are non-negative. We have thus proved the following theorem.

Theorem 2

The function $g(\mu)$ is a convex, quadratic lower bound on the convex function $f(\mu)$.

The following theorem describes conditions for which an optimal solution to (MM) is also efficient.

Theorem 3

Let (x^*, λ^*) be an optimal solution to (MM) where x^* is unique. Then x^* is efficient. Furthermore x^* is efficient in any case if $\lambda^* > 0$.

Proof

To the contrary suppose that x^* is not efficient. Hence there exists x' such that $c_i \cdot x' \geq c_i \cdot x^*$ for $i = 1,2,...,r$ with at least one strict inequality. Hence for any non-negative $\{\lambda_i\}$ one obtains

$$\Sigma \lambda_i c_i \cdot x' \geq \Sigma \lambda_i c_i \cdot x^*, \tag{A-1}$$

so in particular (A-1) holds for λ such that $\lambda F \geq g, \lambda \geq 0$. Furthermore, if $\lambda F \geq g$, $\lambda \geq 0$, then

$$\Sigma \lambda_i c_i \cdot x^* \geq \min_{c \epsilon C''} c \cdot x^*. \tag{A-2}$$

Now from (A-1) and (A-2) it follows that $\min_{c \epsilon C''} c \cdot x'$ is bounded below. Thus let c' be a minimizing vector. It then follows from (A-1) and (A-2) that

$$\min_{c \epsilon C''} c \cdot x' = c' \cdot x' \geq c' \cdot x^* \geq \min_{c \epsilon C''} c \cdot x^*$$

and then x' would also be a maximin solution which contradicts the hypothesis.

Now observe that by construction

$$\lambda^* \hat{C} x^* \geq \lambda^* \hat{C} x \text{ for all feasible } x,$$

so that x^* must be optimal for

$$\begin{aligned} &\max \lambda^* \hat{C} x \\ &Ax \leq b \\ &x \geq 0. \end{aligned}$$

But this means that x^* is efficient since $\lambda^* > 0$, e.g., see [3] Corollary 1.4.

Now consider any efficient point x^* and let Λ be the set of weights for which x^* remains optimal for (e-I). Let λ^* be an optimal solution (one exists since Λ bounded) to

$$\min_{\lambda \epsilon \Lambda} \lambda \hat{C} x^* \qquad \text{LP } (x^*)$$

Define

$$\Lambda^* = \{\lambda \mid \lambda \hat{C} x^* \geq \lambda^* \hat{C} x^*, \Sigma \lambda_i = 1, \lambda_i \geq 0\}.$$

Observe that $\lambda \hat{C} x^* \geq \lambda^* \hat{C} x^*$ determines a supporting hyperplane to Λ.

Theorem 4
The pair (x^*, λ^*) is an optimal solution to

$$\begin{array}{ll} \max & \min \ \lambda \hat{C} x \\ Ax \leq b & \lambda \epsilon \Lambda^* \\ x \geq 0 & \end{array}$$

Furthermore Λ^* is the largest set for which x^* is a maximin solution.

Proof
First observe that

$$\lambda^* \hat{C} x^* \geq \lambda^* \hat{C} x \text{ for all feasible } x$$

since x^* is optimal for any $\lambda \epsilon \Lambda$. Furthermore by construction

$$\lambda^* \hat{C} x^* \leq \lambda \hat{C} x^* \text{ for all } \lambda \epsilon \Lambda^*.$$

Hence (x^*, λ^*) is maximin.
Observe now that $\Lambda \subseteq \Lambda^*$ for let $\lambda' \epsilon \Lambda$.
Then

$$\begin{array}{l} \max \ \lambda' \hat{C} x = \lambda' \hat{C} x^* \\ Ax \leq b \\ x \geq 0 \end{array}$$

since x^* is optimal for (e-I) when $\lambda \epsilon \Lambda$.
Suppose that $\lambda' \notin \Lambda^*$, i.e. $\lambda' \hat{C} x^* < \lambda^* \hat{C} x^*$.
One obtains

$$\min_{\lambda \epsilon \Lambda} \lambda \hat{C} x^* \leq \lambda' \hat{C} x^* < \lambda^* \hat{C} x^*$$

which provides a contradiction since

$$\min_{\lambda \epsilon \Lambda} \lambda \hat{C} x^* = \lambda^* \hat{C} x^*.$$

Hence $\lambda' \epsilon \Lambda^*$ so it follows that $\Lambda \subseteq \Lambda^*$.
Now suppose that (x^*, λ') is a maximin pair relative to a set Λ'.
Hence

$$\begin{array}{ll} \max & \lambda' \hat{C} x = \lambda' \hat{C} x^*, \\ Ax \leq b & \\ x \geq 0 & \end{array}$$

i.e. x^* is optimal for (e-I) when $\lambda = \lambda'$. This means $\lambda' \epsilon \Lambda \subseteq \Lambda^*$ so it follows $\Lambda' \subseteq \Lambda^*$.
In [6] Philip shows that $x^* = (5,0)$ is an efficient solution for the case

$$\hat{C} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

and

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 1 \end{pmatrix}, \quad b = \begin{pmatrix} 5 \\ 7 \\ 10 \end{pmatrix}.$$

A straight forward sensitivity analysis shows that x^* is optimal for (e-I) whenever $\lambda_1 \geq 2\lambda_2$ e.g. $\lambda_1 = 0.8, \lambda_2 = 0.2$.

Vol. 5
1977

It follows that LP(x^*) in this case is

$$\begin{aligned} \min\ & 10\lambda_1 - 5\lambda_2 \\ & \lambda_1 - 2\lambda_2 \geq 0 \\ & \lambda_1 + \lambda_2 = 1 \\ & \lambda_1, \lambda_2 \geq 0 \end{aligned}$$

which yields an optimal solution $\lambda^* = (2/3, 1/3)$. According to Theorem 4 $x^* = (5,0)$ is maximin for (λ_1, λ_2) such that

$$\begin{aligned} & 10\lambda_1 - 5\lambda_2 \geq 5 \\ & \lambda_1 + \lambda_2 = 1 \\ & \lambda_1, \lambda_2 \geq 0 \end{aligned}$$

which yields $\lambda_1 \geq 2/3$. In this case $\Lambda = \Lambda^*$.

Finally let us consider the construction of the lower bound function, $g(\mu)$, for Philip's problem relative to $\lambda_1 = 0.8$, $\lambda_2 = 0.2$. The maximin solution for Philip's problem in the case of total ignorance is $\bar{x} = (5,5)$ with maximum value 5, i.e. $g(1) = 5$. Also

$$\min_{c \in C_1} c \cdot \bar{x} = (0.2, 0.8)\, \hat{C}\bar{x} = 5$$

and

$$\min_{c \in C_2} c \cdot x^* = \min_{i = 1,2} \{c_i \cdot x^*\} = -5$$

Since $g(0) = (0.2, 0.8)\, \hat{C}x^* = 7$, the lower bound function simplifies to

$$g(\mu) = 12\mu^2 - 14\mu + 7.$$

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977. Pergamon Press. Printed in Great Britain

Marketing Segmentation Research— An Application to Bank Services

RL FITTS

Motion Industries Inc., Birmingham, AL 35301, USA

JB MASON

University of Alabama

(Received May 1976; in revised form October 1976)

The research reported here analyses data relating to bank services obtained by a stratified random sampling technique using stepwise linear regression and canonical correlation for the purpose of illustrating the application of segmentation principles to marketing research.

1. 5
77

THE SCOPE of market segmentation research has been limited almost entirely to frequently purchased nondurable consumer goods. A gap exists in segmentation research for consumer durable goods and for services [11], especially commercial bank services. Three bank services have been investigated in some detail: savings accounts [4, 18], bank credit cards [2, 20], and cash dispensing machines [22]. Few replications, however, have been reported concerning usage rates or patterns for any of the three services. Whether specific market segments exist with varying usage rates and identifiable characteristics for various products and services is always a key question for marketers. Previous studies show, however, that a framework does exist within which to study segmentation patterns. Socio-economic, demographic, and attitudinal characteristics have been related statistically to consumption rates [9, 10], brand loyalty [6, 7], and various other dimensions of consumer purchasing behavior [8, 16].

The literature suggests that dimensions of purchasing behavior are limited only by the imagination of the researcher. Furthermore, each dimension may be a practical segmentation criterion. Thus, the critical consideration is whether a dimension of purchasing behavior can be related to different consumption rates and whether, in turn, segments with varying consumption rates can be identified by readily measurable characteristics. The extent to which segmentation research efforts meet these criteria will determine the degree of success of subsequent marketing strategies.

This research was designed to illustrate the application of segmentation principles in marketing research. Specifically by focusing on commercial bank

services, the research was designed to operationalize a generalized market segmentation framework which would, in this research, isolate specific market segments for a package of commercial bank services.

THE DATA BASE

Cross-sectional data were collected from two major Southeastern cities in the United States by use of a stratified random sampling technique. A total of 800 households were interviewed by telephone. The confidence interval was 98% with an allowable standard error of 3%. A total of 768 usable questionnaires was obtained.[1]

Dependent variable—services analyzed

The research was designed to isolate characteristics of users versus nonusers of the following eight commercial bank services: (1) safe deposit boxes; (2) trust services; (3) installment loans; (4) certificates of deposit or Treasury bills; (5) drive-in windows; (6) automatic tellers; (7) overdraft checking; (8) bank credit cards. These services were chosen because previous research has shown that most respondents are familiar with all of these services and that most banks offer all of these services [19] whilst both direct revenue producing and customer attraction/customer maintenance services are represented, thus yielding a cross section of bank services.

The level of use of the services was examined by treating the services as a product group. Predictors were then sought for the use of the group of services.

Independent variable specification

The dependent variables were analyzed in relation to 19 independent variables: three were demographic, seven were socio-economic and nine were attitudinal. The socio-economic and demographic variables were those reported in previous segmentation studies as having meaningful explanatory power [5, 12, 21, 23]. The attitudinal measures represent a relatively new dimension in segmentation research. However, the concept of attitudes as influencing consumption rates is well documented in the literature of consumer behavior [1, 13].

The study hypothesized that as attitudes toward selected commercial bank attributes become increasingly positive, service use will increase. Attitudes toward the attributes were measured by the respondent's degree of satisfaction with those attributes. The attributes analyzed are: (1) loan policy; (2) convenience of location; (3) correctness of statements; (4) adequacy of parking spaces;

[1] The sample was selected from city directories. Preliminary letters were sent to each potential respondent in an effort to reduce the refusal rate. Interviews were conducted both during the day and at night so that both male and female heads of households would be represented in the sample. An oversampling method was employed so that substitution could be made for any sampling unit not responding after three attempts to collect the data.

Vol. 5
1977

(5) adequacy of drive-in facilities; (6) physical appearance; (7) friendliness of personnel; (8) competence of personnel; (9) hours open by bank.

Each respondent's degree of satisfaction was reflected in a 'satisfaction index' which was constructed in the following manner. Each respondent was asked to indicate a degree of satisfaction with each attribute on a modified semantic differential scale. The scale had five points ranging from very satisfied to very dissatisfied. Each point was assigned a weight which assumed a value of five at the favorable extreme of the scale and a value of one at the unfavorable extreme [24]. A 'not applicable' response was also provided and was scored zero [14]. Low respondent scores on each attribute thus reflected less satisfaction with that attribute while high scores reflected greater satisfaction.

The demographic and socio-economic measures used were:

1. *Demographic*: age, sex, marital status.
2. *Socio-economic*: occupation, education, income, length of residence (geographic mobility), number of banks with which household deals, kinds of bank accounts used by household, length status of accounts.

DATA ANALYSIS

Stepwise linear regression and canonical correlation were employed in the analysis of the data.

Model assumptions

The multicollinearity, linearity and homoscedasticity assumptions of the model were initially evaluated.

In evaluating multicollinearity, only one relatively high correlation coefficient was found. Number of different accounts and households having only a checking account were negatively correlated at $r = -0{\cdot}718$. Using Huang's criterion, this degree of correlation was not serious [15].

Three types of plots were used to test for linearity. The first set of plots related the observed value of the dependent package of services variable to each non-dummy independent variable. The second set of plots related the observed value of the package of services variable to the regression residuals. The third set related the predicted value of the package of services variable to the regression residuals. No indications of a non-linear trend were found.

The existence of homoscedasticity was also examined by plotting regression residuals. The predicted value of the package of services variable when plotted against the regression residuals indicated no evidence of heteroscedasticity.

The package of services model

The purpose of the regression analysis was to determine the relationship, if any, between the level of use of a package of bank services and the nineteen

selected socio-economic, demographic, and attitudinal variables. Thus, twenty-seven variables were analyzed by the BMDO2R stepwise regression analysis computer program. Several models were derived in separate analyses in order to develop the one that best fits the sample data. The model presented here represents the 'best' of those derived in terms of (1) the coefficient of determination, (2) the standard error of the regression coefficients, and (3) the standard error of estimate.

A portion of the total regression model is reproduced below in Table 1 for convenience to the reader.[2] The variables are arranged in the order of magnitude of their relative contribution to R^2.

The resulting model includes three attitudinal, five socio-economic, and one demographic variable. These nine variables, in total, accounted for 31% of the variation in the level of use of the package of bank services.

Attitudinal variables

Degree of satisfaction with the adequacy of drive-in windows contributed the greatest amount to the total R^2. It represented 40·3% (12·50/0·31) of the total explanatory power of the model. Because of the positive sign of the coefficient, a high degree of satisfaction with drive-in facilities of a bank apparently is

Vol. 5
1977

TABLE 1. THE PACKAGE OF SERVICES MODEL

Independent variable	Coefficient	Variable type	Contribution to R^2
Constant	−0·7311		
Degree of satisfaction with adequacy of drive-in facilities	0·2385	Attitudinal	0·1250
Number of different accounts	0·3519	Socio-economic	0·0657
Annual household income	0·0019	Socio-economic	0·0344
Degree of satisfaction with loan policies	0·1018	Attitudinal	0·0311
Has savings account only	−0·2727	Socio-economic	0·0204
Length of longest account	0·0565	Socio-economic	0·0105
Education of household head	0·1454	Socio-economic	0·0090
Marital status—single	−0·1868	Demographic	0·0074
Degree of satisfaction with hours open for business	0·1481	Attitudinal	0·0054

$R^2 = 0{\cdot}31$;
$F_{(9,746)}$ for regression model = 37·07, $P < 0{\cdot}0001$;
Standard error of estimate = 0·4994.

[2] All variables in this model have been transformed by their square root (Variable$_t$ = $\sqrt{\text{Variable}_t}$). The interpretation of the results does not change, however, since the same model, differing only in the magnitude of the regression coefficients, was derived with non-transformed data. The R^2, F-ratio, and standard error of estimate in the present model are more favorable than their counterparts in the non-transformed model. Thus, to use the model for predictive purposes, the appropriate transformation should be made.

associated with a high level of use of services at that bank. The second attitudinal variable, satisfaction with loan policies, also entered the model early and contributed an additional 10% to R^2. The final attitudinal variable in the model, satisfaction with the hours open for business, entered last and contributed 1·7% to R^2. In summary, the three attitudinal variables contributed 52·1% of the total explained variance in the level of use of bank services.

Socio-economic variables

Five of the thirteen original socio-economic variables entered this model. The socio-economic variables, in total, contributed 0·14 or 45·2% to the final R^2.

The signs of the regression coefficients for number of different accounts, annual household income, length of time for longest account, and education of household head, are positive and indicate that the level of use of bank services increases, respectively, as:[3]

1. The number of accounts owned increases.
2. The annual household income increases.
3. The length of time of patronage of a particular bank increases.
4. The level of education increases.[4]

Demographic variables

Only one of the original four demographic variables entered the model; moreover, it entered next to last and contributed only 2·4% to the explanatory power of the model. This variable, marital status—single, is a dummy variable and is interpreted as part of a larger variable measuring the marital status of a respondent. Marital status was coded so that its effect on the level of use of bank services is interpreted with respect to the base component, married respondents. The negative coefficient implies that single persons tend to use fewer services than do those who are married.

[3] The interpretation of the relationship between a specific independent variable and the dependent variable by the size and/or sign of the regression coefficient assumes that the other independent variables are held constant.

[4] The fifth socio-economic variable to enter the model, maintains a savings account only, is a dummy variable and requires explanation. This variable was developed using a coded notation which identified the type of accounts owned by the sample respondents. The types of accounts owned were divided into three components—savings account only, checking account only, and both accounts. The coding procedure specifically identified the former two and inferred the latter. Thus, the two identified components are interpreted with reference to the inferred component as a base. Therefore, the negative coefficient signifies that respondents who own only a savings account tend to use fewer bank services than do those who own both checking and a savings account.

Regression analysis summary

The results of this segmentation model compare favorably with other studies reported in the literature. A major contribution, also, is the explanatory ability of another set of independent variables—attitudinal. While the major contributors to R^2 as revealed in the literature have been socio-economic and demographic variables, the present results yielded an almost even split in favor of attitudinal measures. It appears therefore, that more emphasis should be placed on attitudinal measures in segmentation models in developing better predictive and explanatory power.

Canonical correlation analysis was the second method of analysis used in this research. This technique is useful in analyzing the relationship between a set of criterion variables and a set of predictor variables. Thus, the purpose of this portion of the analysis was to search for patterns of use which are associated with certain socio-economic, demographic, and attitudinal characteristics.

Canonical analysis derives two sets of weighting coefficients (a set for the criterion variables and a set for the predictor variables). Thus, two canonical variables are derived—one representing a linear combination of the variables in the criterion set and one representing a linear combination of the variables in the predictor set. The first correlation can be obtained by any pair of linear combinations of these variables.

Vol. 5
1977

Canonical pair 1

Analysis of the criterion variable coefficients indicated that, using a cut-off point of 0·10, three variables were of primary importance (the canonical correlation was 0·61). These were use of drive-in windows, use of loan-services, use of bank credit card. The largest predictor variable coefficients, using a cut-off point of 0·15, were sex of respondent, number of different accounts, satisfaction with adequacy of parking, satisfaction with adequacy of drive-in windows, age of household head, and education of household head. The model for the first canonical pair suggested that users of the combination of drive-in windows, installment loans, and bank credit cards (when contrasted to non-users) tend to: (1) be female; (2) own more bank accounts; (3) have less favorable attitudes toward their bank's parking facilities; (4) have more favorable attitudes toward their bank's drive-in facilities; (5) be younger; (6) have more education.

Canonical pair 2

Although a complete canonical analysis extracts a number of pairs of variates equal to the lesser of the number of criterion or predictor variables, the present analysis yielded two pairs which had canonical correlations over 0·50. Further analysis would have become progressively less meaningful.

The second set of canonical coefficients suggested that non-users of drive-in facilities and users of installment loans, certificates of deposit/treasury bills, and

bank credit cards (when contrasted to users of the former and non-users of the two latter) tend to: (1) have owned a bank account longer; (2) have more favorable attitudes toward their bank's loan policies; (3) have less favorable attitudes toward their bank's drive-in window facilities; (4) have more favorable attitudes toward their bank's business hours; (5) be lower level white collar and upper level and lower level blue collar workers.

DISCUSSION

This research sought to make two contributions to the market segmentation literature: (1) to extend the application of market segmentation models to commercial bank services and (2) to examine and illustrate the importance of selected demographic, socio-economic, and attitudinal variables as they relate to use of bank services.

The regression model developed in this research performed better, in general, than most of those previously reported in the literature. For example, except for an occasional R^2 over 0·30 for a specific product, most have averaged less than 0·20 [7, 8, 21]. In addition, one of the few bank service regression models reported in the literature examined users versus non-users of bank credit cards and accounted for 6% of the variation in the level of use ($R^2 = 0{\cdot}06$) with social class, income, and life cycle variables. Thus, the present model performed considerably better than previous models.

The predictive ability of the models presented was somewhat disappointing but not totally surprising. The results are comparable to other contemporary market segmentation research findings. The results are encouraging, however, in that attitudinal variables were significant in the models. Moreover, high correlations and coefficients of determination cannot be expected when either the criterion or predictor variables are dichotomous or scale values, which, up to this time, is often the case in marketing studies. Further research on refining our ability to segment markets effectively is clearly needed.

REFERENCES

1. ACHENBAUM A (1966) Knowledge is a thing called measurement. *Attitude Research at Sea* (Eds. ADLER L and CRESPI I), pp. 111–126. American Marketing Association, Chicago.
2. AWH R and WATERS D (1974) A discriminant analysis of economic, demographic, and attitudinal characteristics of bank charge card holders. *J. Fin.* **29**, 973–980.
3. BELL JE (1969) Mobiles—a neglected market segment. *J. Marketing* **33**, 37–44.
4. CLAWSON CJ (1974) Fitting branch locations, performance standards, and marketing strategies to local conditions. *J. Marketing* **38**, 8–14.
5. CUNNINGHAM R (1962) Customer loyalty to store brands, *Harv. Bus. Rev.* **39**, 127–137.
6. FRANK RE and BOYD HW (1965) Are private-brand prone grocery customers really different? *J. Advertising Res.* **5**, 27–35.
7. FRANK RE, DOUGLAS SP and POLLI RE (1968) Household correlates of brand loyalty for grocery products. *Jl Bus.* **41**, 237–245.

8. FRANK RE, GREEN PE and SIEBER HF (1967) Household correlates of purchase price for grocery products. *J. Marketing Res.* **4**, 381–384.
9. FRANK RE, MASSY WF and BOYD HW (1967) Correlates of grocery product consumption rates. *J. Marketing Res.* **4**, 184–190.
10. FRANK RE, MASSY WF and LODAHL TM (1969) Purchasing behavior and Personal Attributes, *J. Advertising Res.* **9**, 15–24.
11. FRANK RE, MASSY WF and WIND Y (1972) *Market Segmentation.* Prentice-Hall, Englewood Cliffs.
12. GOLDSTEIN S (1968) The aged segment of the market: 1950–1960. *J. Marketing* **32**, 62–68.
13. HELLER HE (1968) Defining Target Markets by their Attitude Profile, *Attitude Research on the Rocks* (Eds. ADLER L and CRESPI I), pp. 45–57. American Marketing Association, Chicago.
14. HUGHES GD (1967) Selecting scales to measure attitude changes. *J. Marketing Res.* **1**, 85–87.
15. HUANG DS (1970) *Regression and Econometric Methods.* John Wiley, New York.
16. KINNEAR TC, TAYLOR JR and AHMED SA (1974) Ecologically concerned consumers—who are they? *J. Marketing* **38**, 20–24.
17. KOPONEN A (1960) Personality characteristics of purchasers. *J. Advertising Res.* **1**(1) 6–12.
18. MARTIN P (1967) Savings and loans in new submarkets—search for predictive factors. *J. Marketing Res.* **4**, 184–190.
19. MASON JB and MAYER ML (1974) Analyzing factors that affect banking patronage. *Bank Admin* **50**.
20. MATHEWS HL and SLOCUM JW (1972) Correlates of commercial bank credit card use. *J. Bank Res.* **2**, 20–27.
21. PESSEMIER EA, BURGER PC and TIGERT DJ (1967) Can new product buyers be identified? *J. Marketing Res.* **4**, 349–355.
22. POOL AA (1974) Applications of discriminant analysis in formulation of promotional strategy for cash dispensing machines. *J. Bank Res.* **5**, 13–19.
23. WELLS WD and GUBAR G (1966) The life cycle concept in marketing research. *J. Marketing Res.* **3**, 355–363.
24. WIND Y (1968) Integrating attitude measures in a study of industrial buying behavior. *Attitude Research on the Rocks* (Eds. ADLER L and IRVING I) pp. 97–125. American Marketing Association, Chicago.

ADDRESS FOR CORRESPONDENCE: *Dr RL Fitts, Director of Personnel and Marketing, Motion Industries Inc., PO Box* 1477, *Birmingham, AL* 35301, *USA*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977. Pergamon Press. Printed in Great Britain

Quantitative Models for Optimal Rest Period Scheduling

GL GENTZLER

Bell Telephone Laboratories, Holmdel, USA

TM KHALIL

The University of Miami

BD SIVAZLIAN

The University of Florida

(*Received February 1976; in revised form October 1976*)

ol. 5
977

Mathematical and simulation models studying work–rest schedules in a production process are developed to arrive at an optimal rest policy to maximize work output per unit time. A CSMP simulation study is used to test the sensitivity of the results for different specific cases.

INTRODUCTION

A DECREMENT in human performance is usually associated with continued work. In industry, fatigue is the chief factor in limiting production, because it decreases efficiency [10]. This causes a corresponding decrease in the rate of production or work output [1].

The standard method of reducing fatigue is the introduction of rest periods. It has been known that daily productivity is increased by the use of rest periods [1]. Studies have shown that after such rest periods, efficiency shifts back to the upper portion of its range [12, 15, 16], therefore, increasing overall productivity of an individual.

Frequent rest periods will allow a high level of efficiency to be maintained, but at the cost of increased idle time during the rest periods. Few rest periods reduce the amount of idle time, but then most of the time the individual will be operating at low levels of efficiency. An optimal rest schedule could be that for which work output over time is maximized.

Research on work–rest schedules has traditionally concentrated on the experimental approach to determine optimal work–rest schedules for specific tasks and under specified environmental conditions [3, 4, 7, 14, 15, 16].

Only a few attempts have been made to use quantitative methods to model man–machine systems and to optimize scheduling of rest periods [8, 9, 13]. Eilon [6] modeled the influence of fatigue and rest periods on production rate. For this model, production over a finite time interval of length T is to be maximized by initiating a single rest period of length δ at the beginning of the interval, where the production rate is r. The remainder of the interval is spent working. The assumption is made that production rate decreases as activity proceeds. Also, the production rate function shifts upward by $a\delta$ after a rest period, while it declines at the rate b if work continues. Under these conditions, maximum production corresponds to

$$\delta^* = \frac{(a+b)\,T - r}{2a+b}.$$

WORK–REST MODEL

Vol. 5
1977

The process considered deals with a continuous series of work–rest cycles over time. The nature of the task is such that a rest period of nominal length results in efficiency returning to the unfatigued state. In turn, the effect of one work period upon another is considered negligible. This is a reasonable assumption for tasks requiring a moderate degree of physical or mental exertion. Also, no external stresses are considered.

The basic component of the model is the *work rate function*, which defines the work rate, or performance level, from the end of one rest period to the beginning of the next rest period. Typically the work rate would start at a level representing the individual's work rate in the 'unfatigued state'. As work proceeds, it may increase at first due to a 'warm-up' period, or it may start to decrease from the beginning. Once it starts to decrease, fatigue will cause it to do so, until the next rest period, thus unimodality holds. This function has been practically identified by several researchers studying repetitive work in industry [5, 11].

With the above assumptions, the production process can be modeled as a series of identical work–rest cycles, as shown in Fig. 1. For a work–rest cycle let

x = the length of time of a work period;

y = the elapsed time since the beginning of the present work period;

θ = the length of a rest period; a constant long enough to ensure that the work rate of the next work period starts at $v(0)$, the unfatigued state;

$\phi(t)$ = the amount of work output produced up to time t;

$v(y)$ = the instantaneous rate of work or the work rate function which starts at $y = 0$ for each work period; $v(y) = \phi'(y)$.

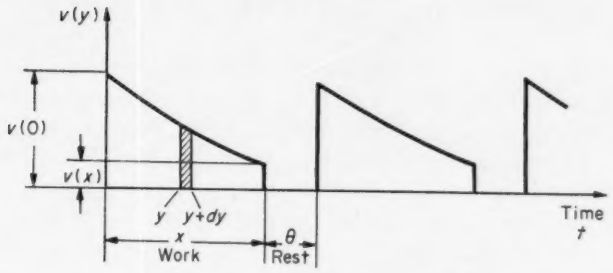


FIG. 1. *Behavior of the work rate function v(y).*

For a given production process, or task, $v(y)$ and θ would be specified. Therefore, the decision variable used in determining a suitable work–rest policy is x.

THE OPTIMAL POLICY

The optimal policy will define x^*, where x^* is the value of x that yields the maximum work output per unit time. Since the process is cyclic, x^* can be found by maximizing the work output per unit time of a single cycle, which is

$$W(x) = \frac{\phi(x)}{x + \theta}.$$

The objective function is then

$$\max_x W(x).$$

Setting $W'(x) = 0$ yields

$$\phi'(x) = \frac{\phi(x)}{(x + \theta)}$$

or

$$v(x) = W(x) \quad \text{at} \quad x = x^*.$$

This states that the optimal time to introduce a rest period is when the average work rate of a cycle is equal to the instantaneous work rate. The sufficiency condition for maximum is $W''(x) < 0$. At x^*, this reduces to $\phi''(x^*) < 0$. Therefore, a sufficient condition for an interior maximum is that $v'(x) < 0$, that is, $v(x)$ be a strictly decreasing function of x. However, it is still possible for a unique local interior maximum to exist in general because $W(x)$ will usually be unimodal with an interior maximum.

COMPARISON WITH EILON'S RESULT

In the context of this discussion, Eilon's model deals with the case where only a limited amount of production time is left. Thus, a rest period long enough to allow full recovery to the unfatigued state is not dictated. As expected, applying the necessary condition $v(x) = W(x)$ does not ensure optimality for the process addressed by Eilon. That is, the condition r equals the average work rate over T is not necessarily satisfied when the length of the rest period is δ^*, the value yielding maximum production. (The solution is δ equals δ^* plus or minus a term not always reducing to zero). This shows that the relationship between rest period and recovery is central to the optimal policy.

Examples

Using the results of the last section, x^* was determined for two types of work rate functions—linear and exponential. For each case it is assumed that $v(y) > 0$ in the range of interest. First, $\phi(x)$ is calculated by taking the integral of $v(y)$ from 0 to x. Then, $\phi(x)$ and $v(x)$ are substituted into the equation of the necessary condition, which is solved for x^*. The results obtained for the two cases are:

(i) *Linear*, with $v(x) = b - ax$ (a, b positive constants).

The necessary condition yields a quadratic equation, whose positive root is

$$x^* = -\theta + \sqrt{\left(\theta^2 + \frac{2b\theta}{a}\right)}.$$

(ii) *Exponential*, with $v(x) = e^{-ax}$.

The necessary condition yields the expression

$$x^* + \theta + \frac{1}{a} = \frac{1}{a} e^{ax^*}.$$

where upon expanding e^{ax^*} as a power series results in the approximation

$$x^* \simeq \sqrt{\frac{2\theta}{a}}.$$

The sufficiency condition is satisfied in both cases.

SIMULATION STUDY

The purpose of this simulation study is to give numerical examples of the production process, to serve as a sensitivity analysis of the effects of different policies on work output, and to present a simulation technique that can be extended to solve problems not suitable for analytic methods. In this study, a linear and an exponential work rate function is used; the parameters of each are arbitrarily set so that $x^* = 60$ min for $\theta = 10$ min in the theoretical formulas

Vol. 5
1977

($b/a = 240$ for the linear case and $a = 1/180$ for the exponential case). A finite time horizon of 240 min is used, representing morning to noon. This does not correspond to an integer number of cycles for x^*, so it should give an indication of how this 'poor end condition' affects the optimal policy. For each work rate function the x^* policy is compared with that of using from zero to five evenly spaced rest periods. Evenly spaced means that the rest periods are placed so that all the work periods have the same length, with a complete work period at the beginning and at the end of the time horizon. The x^* policy will thus result in a partial work period at the end of the horizon, but there will be three full rest periods.

The CSMP computer program used to simulate the production process is given by the block diagram of Fig. 2. In this model the work output W is the integral of the work rate V. The value of V is equal to the work rate function during work periods and zero during rest periods. The function REST, a pulse function, controls the value of V and resets the work rate function back to $v(0)$ at the end of each rest period. The impulse function P defines the rest period schedule. The units of work are arbitrarily chosen, so that when $v(y)$ equals one, one unit of work is produced per minute.

It was found that the x^* policy was at least as good as any other, even though there was not an integer number of cycles. However, the policies for two and three evenly spaced rests resulted in total work output as large as the x^* policy. This indicates some insensitivity of the work output for rest policies for these cases.

Although it is possible to derive the x^* policy under the more stringent assumption of a finite time horizon, the derivation will not be presented here. It suffices, however, to say that under this latest assumption, the optimal policy is to have evenly spaced rest-periods, with the number of rest periods chosen so

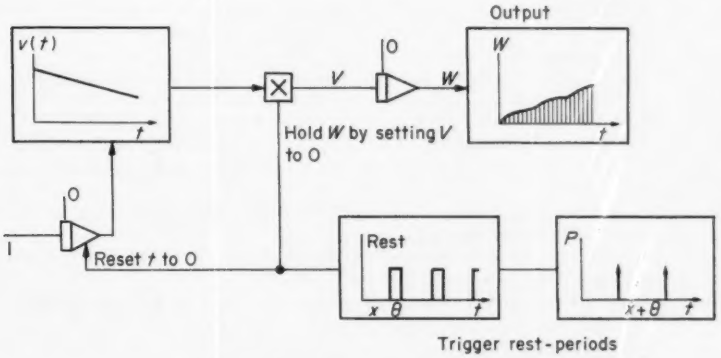


FIG. 2. *CSMP block diagram of the work–rest model.*

that x is close to that value of x^* dictated by the infinite horizon problem. The technique of dynamic programming [2] can be successfully used to model this situation and arrive at optimal policies.

CONCLUSIONS

The normative models for optimal rest period scheduling discussed in this paper are useful in evaluating organizational policies and in determining optimal requirements for achieving maximum productivity. Attention should however be drawn to the fact that the models are purely mechanistic. Practical validation using samples of real workers for different types of situations would determine the applicability of the models to particular cases.

The present basic work–rest models can be modified to include a stochastic environment, and the models can be expanded to represent more complex situations. Also, the simulation model can be further extended to incorporate the effects of factors other than fatigue on work output.

REFERENCES

1. Barnes RM (1968) *Motion and Time Study*, pp. 570–579. John Wiley, New York.
2. Bellman R (1957) *Dynamic Programming*, Princetown University Press, Princeton, NJ.
3. Colquhoun WP (1959) 'The effect of a short rest/pause on inspection efficiency.' *Ergonomics*, **2**, 367–372.
4. Colquhoun WP, Black MJF and Edwards RS (1969) 'Experimental studies of shift work III: stabilized 12-hour shift systems. *Ergonomics* **12**, 865–882.
5. Dudley NA (1958) Output pattern in repetitive tasks. *Instn Prod. Engrs J.* **37**, 187.
6. Eilon S (1964) On a mechanistic approach to fatigue and rest periods. *Int. J. Prod. Res.* **3**, 327–332.
7. Khalil TM and Ayoub MM (1976) Work scheduling under normal and prolonged vibration. *AIHAJ* 174–182.
8. Khalil TM and Gentzler G (1972) Hybrid computer simulation of a man–machine system. *Proceedings of the International Conference on Cybernetics Society*, Washington DC, 141–142.
9. Khalil TM and Valisalo PE (1973) A hybrid computer approach for solving stochastic man–machine systems. *Proceedings of the 7th International Analog Computation Meeting, Prague, Czechoslovakia*, 69.
10. McFarland RA (1971) Understanding fatigue in modern life. *Ergonomics* **14**, 1–10.
11. Murrell KFH (1965) *Human Performance in Industry*. Reinhold, New York.
12. Niebel BW (1972) *Motion and Time Study* pp. 104–107. Irwin, Homewood, IL.
13. Proctor CL, Khalil TM and Sueve CC (1971) A quantitative approach to performance evaluation of man–machine systems having a stochastic environment. *Int. J. Man-Machine Studies* **3**, 127–140.
14. Ramsey JD, Halcomb CG and Mortagy AK (1974) Self-determined work/rest cycles in hot environment. *Int. J. Prod. Res.* **12**, 623–631.
15. Rhomert W (1973) Problems in determining allowances; part I: use of modern methods to evaluate strain in static muscular work. *Appl. Ergonomics* **4** (2), 91–95.
16. Rhomert W (1973) Problems of determination of rest allowances; part 2: determining rest allowances in different human tasks. *Appl. Ergonomics* **4** (3), 158–162.

ADDRESS FOR CORRESPONDENCE: *Professor BD Sivazlian, University of Florida, College of Engineering, Gainesville, Florida* 32611, *USA*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 2, 1977. Pergamon Press. Printed in Great Britain

International Register of Theses and Dissertations

BRETON L (1973) *French: Une simulation de trafic maritime dans l'optique de la Dynamique Industrielle.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

CASIEZ S (1973) *French: Etude des caractéristiques d'usage des véhicules récréatifs.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

COLASSE P (1972) *French: Limites des modèles de simulation générale d'entreprise.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

DIF J (1973) *French: L'utilisation des obligations convertibles dans le financement des entreprises françaises.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

HEDTMANN B (1973) *French: La demande de monnaie par les firmes.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

HITZBLECK H (1972) *Die Bedeutung des Fisches für die Ernährungswirtschaft Mitteleuropas in vorindustrieller Zeit unter besonderer berücksichtigung Niedersachsens.* PhD Thesis, George August-Universität Göttingen. (Available at: Neidersachsichen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prinzenstrasse 1, Germany)

KABREDE HJ (1972) *Zur Theorie der Mehrprodukt—und Mehrstufenunternehmung.* PhD Thesis, George-August-Universität Göttingen. (Available at: Niedersachsichen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prinzenstrasse 1, Germany)

LETZ WB (1972) *'Werbung und Verkündigung'—Zur Ubertragung sozialpsychologischer Denkmodelle in das religiöse Feld unter dem Gesichtspunkt der 'Sozialen Steuerung'.* PhD Thesis, George-August-Universität Göttingen. (Available at: Niedersachsichen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prinzenstrasse 1, Germany)

LIENAU R (1972) *Ausserungen des Landsmannschaftlichen Bewusstseins.* PhD Thesis, George-August-Universität Göttingen. (Available at: Niedersachsichen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prinzenstrasse 1, Germany)

MAI D (1972) *Untersuchung informationstheoretischer Ansätze bei der absatzwirtschaftlichen Aktionsanalyse.* PhD Thesis, George-August-Universität Göttingen. (Available at: Niedersachsichen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prinzenstrasse 1, Germany)

MOHLMANN H (1972) *Die Entwicklung eines integrierten Gesamtplanungssystems—unter besonderer Berücksichtigung eines Betriebes der Gemeinnützigen Wohnungswirtschaft.* PhD Thesis, George-August-Universität Göttingen. (Available at: Niedersachsichen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prinzenstrasse 1, Germany)

REBOUH B (1973) *French: Le seuil de rentabilité: Instrument d'analyse de gestion.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

SCHIEB PA (1973) *French: L'efficacité differentielle de la publicité—Extension–Adaptation–Invention.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

THIEBLEMONT R (1973) *French: Etude de la gestion de Trésorerie de la ville de Sherbrooke.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

TORNATO P (1973) *French: Etude des facteurs déterminants de l'image d'une grande société: le cas de Bell Canada.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

VEYS P (1972) *French: Le marketing du ski; une application pratique.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

The following are obtainable at the Library, Cranfield Institute of Technology, Cranfield, Bedford MK43 0AL, UK:

BAJWA SH (1973) The Application of a Numerically Controlled Machine Tool to the Inspection of Components. MSc Thesis, Cranfield Institute of Technology.

BARGHUTHI A (1973) Aspects of Process Control in a Canned Food Plant. MSc Thesis, Cranfield Institute of Technology.

CALLOW JW (1970) An Investigation into Incentives with Special Reference to the Effectiveness of Payment by Results in the Field of Specialised Engineering Products. MSc Thesis, Cranfield Institute of Technology.

CHATTOPADHAY D (1970) Capacity Utilization in a Large Switch-Gear Manufacturing Organisation. MSc Thesis, Cranfield Institute of Technology.

COOMBES DJ (1970) An Investigation into Factors Affecting the Optimum Utilisation of Numerically Controlled Machine Centres in Aircraft Component Manufacture. MSc Thesis, Cranfield Institute of Technology.

CORDRAN EC (1971) An Investigation into the Performance of New Numerically Controlled Machine Tools as a Basis for Increasing the Output in a Turbogenerator Machine Shop. MSc Thesis, Cranfield Institute of Technology.

DHAWAN R (1971) Methods-Time Measurement in Torque Application on Bolts. MSc Thesis, Cranfield Institute of Technology.

DUGGAL SC (1971) Analysis and Design of the Stock Control for a Small Manufacturing Company. MSc Thesis, Cranfield Institute of Technology.

DUNNE WT (1971) A Practical Production Control System for Manufacture to Customers' Order. MSc Thesis, Cranfield Institute of Technology.

EDE MJ (1973) An Application of Group Technology to Batch Manufacture. MSc Thesis, Cranfield Institute of Technology.

FAULDER KA (1973) Estimation of the Possible Increase in Natural Frequency of a Planning Machine Using Optimum Design Techniques. MSc Thesis, Cranfield Institute of Technology.

FLINT D (1973) An Improved Method of Production and Stock Control in a Batch Manufacturing Organisation. MSc Thesis, Cranfield Institute of Technology.

GEE DM (1971) Wage and Salary Structure Evaluation (Chrysler UK Ltd). MBA Thesis, Cranfield Institute of Technology.

GRINDELL RJ (1971) The Cost of One-Man Bus Operation. MBA Thesis, Cranfield Institute of Technology.

HAYS MS (1971) A Feasibility Study of the Introduction of a New Product into the Contract Floorcovering Market. MBA Thesis, Cranfield Institute of Technology.

JONGEBREUR RS (1971) An Analysis of a Sub-Contract Machining Company. MBA Thesis, Cranfield Institute of Technology.

LAKE WJB (1971) Practical and Economic Considerations of a Proposed Conversion from the Tropenas Converter Melting Plant to Electric Arc Furnaces at Lloyds (Darlington) Ltd. MBA Thesis, Cranfield Institute of Technology.

Vol. 5
1977

LANE WR (1971) A Survey of some Aspects of Services to the Deaf and Hard of Hearing. Possible Developments and Opportunities for Commercial Hearing Aid Companies. MBA Thesis, Cranfield Institute of Technology.

LIEN SKH (1972) The Potential for Commercial and Industrial Development in the Portsmouth–Southampton Area. MBA Thesis, Cranfield Institute of Technology.

LINDSEY N (1972) A Corporate Strategy for a Hardware and Electrical Products Wholesaler. MBA Thesis, Cranfield Institute of Technology.

LUI WF (1972) Investment Analysis of Ten Tin Mining Companies Operating in Malaysia. MBA Thesis, Cranfield Institute of Technology.

LUMSDON CA (1972) The Individual in the Organisation. MBA Thesis, Cranfield Institute of Technology.

MCFIGGINS DM (1970) A Survey of a Company in the Process of an Organisational and Technological Change to Reveal Possibilities and Opportunities. MBA Thesis, Cranfield Institute of Technology.

MAKIN P (1970) Assessment of Business Potential of a Modular System for Real Time Data Handling Equipment. MBA Thesis, Cranfield Institute of Technology.

MAJUMBAR AK (1972) An Investigation into the Analysis of Profitability of Small Manufacturing Firms. MSc Thesis, Cranfield Institute of Technology.

MANDOUVALOS E (1971) A Survey in Corporate Planning. MBA Thesis, Cranfield Institute of Technology.

MARI G (1972) The Problem of Marketing Italian Speciality Foods in Great Britain. MBA Thesis, Cranfield Institute of Technology.

MARSH PJ (1972) An Investigation into some Factors Affecting the Success of Drive-In Wimpy Bars. MBA Thesis, Cranfield Institute of Technology.

MARTIN WG (1970) The Industrial Buying Function in the Electronic Components Industry. MBA Thesis, Cranfield Institute of Technology.

MATHER CT (1970) A Study of the Changing Market Structure, Needs and Opportunities for a Major Marketing Division. MBA Thesis, Cranfield Institute of Technology.

MEHTA CB (1972) Effects of a New Manufacturing Line on Other Work-Centres. MSc Thesis, Cranfield Institute of Technology.

MOIR RP (1973) The Design and Development of a Rotary Index Table Drive System Employing Linear Motors. MSc Thesis, Cranfield Institute of Technology.

PADMANABHAN V (1970) An Investigation into the Methods and Problems Associated with Planning for Aircraft Production in India. MSc Thesis, Cranfield Institute of Technology.

PEGG PR (1971) Volumetric Calibration of Machine Tools with Particular Reference to the Effect of Heat. MSc Thesis, Cranfield Institute of Technology.

PREECE M (1973) An Alternative Manufacturing System to Batch Production in a Machine Shop. MSc Thesis, Cranfield Institute of Technology.

QUICK NJ (1971) The Implications of Possible Improvements in the Utilisation of Coiled Steel in Chrysler United Kingdom Limited (Joint Thesis with MN HILDERBRANDT). MBA Thesis, Cranfield Institute of Technology.

REID CDD (1972) An Investigation into Self-Checking Number Systems, particularly Applied to a Computerised Hospital Pathology Laboratory. MSc Thesis, Cranfield Institute of Technology.

STOTLAND JA (1972) Seeking and Acquiring a Small Business—A Practical Case. MBA Thesis, Cranfield Institute of Technology.

SWEETING EM (1972) An Investigation of the Fork Lift Truck Market. MBA Thesis, Cranfield Institute of Technology.

TAYLOR MH (1972) The Economics of Residential Housing in London. MBA Thesis, Cranfield Institute of Technology.

THOMAS LM (1972) Malta Industries Ltd. MBA Thesis, Cranfield Institute of Technology.

VORLEY G (1972) Development and Application of a Computer Controlled Machine for Manufacturing Irregular Profiles. MSc Thesis, Cranfield Institute of Technology.

WALKER LJ (1972) The Effects of Increased Press Output on the Post-Press Operations at an Aluminium Extrusion Plant. MSc Thesis, Cranfield Institute of Technology.

WAPPAT K (1973) An Investigation into the Performance of a Machine Shop and Subsequent Proposals for a Computerised Shop Loading and Scheduling System. MSc Thesis, Cranfield Institute of Technology.

WARREN JF (1972) Forecasting Brand-Share. MSc Thesis, Cranfield Institute of Technology.

WATKIN JG (1971) The Development of a Production Control System for a Design/Prototype Orientated Establishment. MSc Thesis, Cranfield Institute of Technology.

WEST AN (1971) An Investigation into the Overall Effect of Tolerances on Component and Assembly Costs. MSc Thesis, Cranfield Institute of Technology.

WIJESEKARA JB (1972) Analysis of Semi-Automatic Production Systems in the Manufacture of Cement. MSc Thesis, Cranfield Institute of Technology.

WILCOCK GA (1971) A Systematic Approach to Depot Siting Problems with Reference to a Manufacturer of Materials Handling Equipment. MSc Thesis, Cranfield Institute of Technology.

WOOD GM (1972) The Total System Approach to Asphalt Paving. MSc Thesis, Cranfield Institute of Technology.

The following are obtainable at the MIT Microreproduction Laboratory, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02139, U.S.A.:

ANDERSON GW (1973) The External Analysis of Corporate Strategy—Can it be Meaningful? SM Thesis, MIT, Cambridge, MA.

BEAN WH (1971) Specialization of Communication Channels in R&D Laboratories. SM Thesis, MIT, Cambridge, MA.

BHUSHAN AK (1971) Selection and Evaluation of Minicomputers. SM Thesis, MIT, Cambridge, MA.

BISHOP KJ (1971) Strategies for the Agricultural Chemical Industry. SM Thesis, MIT, Cambridge, MA.

BLOOMSBURGH P (1971) Man–Machine Decision Systems: A Design Methodology and an Implemented System for the M.I.T. Capital Budgeting Process. SM Thesis, MIT, Cambridge, MA.

BOWMAN PB (1973) A Decision Oriented Management Information System in a Naval Shipyard. SM Thesis, MIT, Cambridge, MA.

BRODSKY G (1972) A Simulation Model for Corporate Debt Capacity. SM Thesis, MIT, Cambridge, MA.

BROOKS MS (1973) Technology Transfer and the Existence of Exploitable Ideas Among the M.I.T. Faculty. SM Thesis, MIT, Cambridge, MA.

BROSIO Sergio (1973) Valuation Methods and the Use of Computers in the Acquisition Analysis of Privately Owned Firms. SM Thesis, MIT, Cambridge, MA.

BROWN JL (1970) Some Problem-Solving Procedures for the Two-Dimensional Cutting Stock Problem. SM Thesis, MIT, Cambridge, MA.

BULLOCK WB (1970) Management Science in Decision-Making: Case Studies of Seven Selected Companies. SM Thesis, MIT, Cambridge, MA.

BURGER AF (1970) An Interactive Model to Aid in Cost Planning. SM Thesis, MIT, Cambridge, MA.

CHAPIROT GA (1971) The Allocation of Corporate Goals to the Operating Divisions of a Company. SM Thesis, MIT, Cambridge, MA.

CUSICK PV, JR (1971) Design and Implementation of a Management Information System for a Firm in the Construction Industry—A Case Study. SM Thesis, MIT, Cambridge, MA.

DAHANUKAR PD (1973) Offshore Manufacture of Electronic Components by US Electronic Firms—A Case Study. SM Thesis, MIT, Cambridge, MA.

DEBROUWER OJ (1973) The International Finance Corporations—A Member of the World Bank Group. SM Thesis, MIT, Cambridge, MA.

DESHIELDS O (1973) Economic Development Potential via the United States Agency for International Development, the United States Direct Private Investments, and the Japanese Trading Companies. SM Thesis, MIT, Cambridge, MA.

DILL NE (1973) Trends in the Foodservice Industry. SM Thesis, MIT, Cambridge, MA.

ELKUS WS (1973) The Information Content of New Security Issues. SM Thesis, MIT, Cambridge, MA.

FAIRAIZL AF (1973) A Model of Interpersonal Dynamics Within the Community Health Team. SM Thesis, MIT, Cambridge, MA.

FARRAR D (1973) A Model for the Determination of Optimal Electric Generating System Expansion Patterns. SM Thesis, MIT, Cambridge, MA.

FAURE GF (1973) A Study of the Factors Affecting Mutual Funds Net Sales. SM Thesis, MIT, Cambridge, MA.

FINDLEY JJ (1973) Is Prepaid Dental Care the Solution? SM Thesis, MIT, Cambridge, MA.

GEIGER DR (1972) Implementing Aggregate Production Planning. SM Thesis, MIT, Cambridge, MA.

GREENWOOD JM (1972) Contracts as an Incentive to Product Durability, Reliability and Maintainability. SM Thesis, MIT, Cambridge, MA.

GUINNESS TWN (1971) Development of a Framework to Aid the Design of Investment Management Information and Decision Systems. SM Thesis, MIT, Cambridge, MA.

HARMON WR (1971) A Comparative Analysis of the Employment Progress of Disadvantaged Workers. SM Thesis, MIT, Cambridge, MA.

HELMS JD (1971) Communications between Diverse Technical Groups. SM Thesis, MIT, Cambridge, MA.

JERABEK FJ (1971) The Role of the Consumer in Comprehensive Health Planning. SM Thesis, MIT, Cambridge, MA.

JUND LM, JR (1971) A Model of Systematic Fluctuations in a Securities Market. SM Thesis, MIT, Cambridge, MA.

LAFRANCHI LA (1971) The Characteristics of Technological Gatekeepers. SM Thesis, MIT, Cambridge, MA.

LANE JM (1971) Utilization and Benefits of Extended Care Facilities in Massachusetts. SM Thesis, MIT, Cambridge, MA.

MARKOWITZ EA (1971) Causative Factors of Student Disruption and Alternative Methods of Their Resolution. SM Thesis, MIT, Cambridge, MA.

MATHIASEN KJ (1971) M.I.T.'s 'Housing Program in Cambridge'—A Management Decision. SM Thesis, MIT, Cambridge, MA.

MAWSON A (1971) A Study of Divestiture. SM Thesis, MIT, Cambridge, MA.

MENDELSOHN SM (1971) Corporate Planning and Control and its Impact of R&D Effectiveness. SM Thesis, MIT, Cambridge, MA.

PASNIK A (1973) Manpower Planning—An Organisation Development Approach. SM Thesis, MIT, Cambridge, MA.

PAYET-GASPARD PA (1973) An Economic Model of the Aluminium Industry. SM Thesis, MIT, Cambridge, MA.

POOLE DV (1973) Behavioral Sciences/Organization Development within the Health Care System—The Doctor as Client. SM Thesis, MIT, Cambridge, MA.

ROSSOLIMO AN (1973) The Relationship Between the Price Differentials Between National and Private Brands of Supermarket Products and the Advertising-to-Sales Ratios of the National Brands.

Information for inclusion in this Section, presented in the appropriate form, should be sent to the Editor of Omega.

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

EDITORIAL

How Do You Feel About Your Work?

HOW DO you know whether people are happy in their jobs? What they think about the organization in which they work, their colleagues, their bosses? What changes they would like to see in the work environment? What irritants they regard as contributing to a sense of dissatisfaction with their jobs?

The answer seems simple enough: you ask them. Admittedly, there are other ways with which you can try to draw conclusions about the level of job dissatisfaction, such as from measures of absenteeism, labour turnover, slovenly work, untidy workplace, poor quality of production, strikes and stoppages, poor attendance at social functions; and similarly, various positive measures as indicators of job satisfaction can be devised. But all these measures relate to a host of factors other than to whether an employee is happy with his job or not, and it would be far too simplistic to suggest that a change in any one measure is clear proof of a change in the employee's frame of mind. As for the proposition that you should not draw conclusions from any single measure but use several for this purpose, the question arises of how to combine them and what weighting system should be adopted if any attempt is made at quantifying the relationship between a combined yardstick and the level of job satisfaction.

As all such measures may be regarded as mere indicators in our quest for answers to the questions posed above, and as these measures cannot in any sense be acknowledged as necessary or even sufficient conditions from which firm conclusions may be drawn, it follows that a methodology that relies on direct probing rather than on indirect inferences would be thought of as superior, both in terms of efficiency and with respect to the quality of the results that describe the attitudes of the employees. Hence, the simple answer: you ask them. They should know. After all, they experience the work environment, they have to perform their jobs under various constraints, they have to comply with rules and procedures, they encounter irritation and frustration. Who can better tell what effects these factors have than the people who actually experience them?

It is this simple answer to what appears to be a simple methodological question that is at the root of a vast industry, the industry of attitude surveys. (A distinction can be made, as pointed out by Harper and Kynaston-Reeves in their paper published in this issue, between attitude surveys and opinion surveys,

the latter being expressions of beliefs about actual or future facts and events rather than about people's preferences as to particular events and how they should be shaped; however, it is doubtful whether in practice such a distinction is significant, and even the best designed opinion surveys are not immune from the influences of attitudes.) "An assumption of attitude theory is that, all things being equal, the holding of a certain attitude will lead to behaviour that accords with it", say Harper and Kynaston-Reeves, but they are wise to add the qualification: "In real life of course all things are not equal and many factors intervene to prevent behaviour being determined exclusively by a particular attitude a person holds".

Now, the first questions that need to be asked are "Who wants to know and why?" These may seem naive or even irrelevant questions: if the attitude survey is a measuring device, a means by which factual information about workers' attitudes and morale is to be determined, what does it matter "Who wants to know and why?" It may be argued that such questions never arise when an investigator records the air temperature in a given locality or the level of cash balance in a bank account at a given moment in time; his measurements remain valid irrespective of who will use them and for whatever purpose, or at least so it would seem. But in an industrial environment it is difficult to conceive of measurements carried out in a vacuum; they have a purpose, they become the basis for analysis and decision, they need to relate to control mechanisms which adjust conditions and operational parameters in the work system. And consequently the manner in which the measurements are carried out becomes relevant and needs to be specified: the precise place and time at which the temperature or the cash balance are recorded could become important, and so the needs of the user of the information—the decision maker—have to be taken into account in the execution of the measurement activity.

Vol. 5
1977

Lest someone thinks that this is just an obscure academic quibble, let me point out the practical consequences of the proposition that "Who wants to know and why?" is important, which are simply that the design of the attitude survey in a company may well be affected by whether it is to serve the needs of management, the needs of the workforce, or the needs of the unions, to mention but a few of the possible interested parties. In short the survey has a client, the client may have one or several purposes in mind, and what information—and the manner by which it is obtained—is sought by one client may not suit the needs of another user of the said survey.

Is it not possible, you may ask, to serve several clients simultaneously? After all, a temperature is a temperature, so is a bank balance; if two users are interested in the prevailing temperature at different spots, or in the cash level at different points in time, all they need to do is specify their requirements, and the investigator will then repeat his measurements as often as is necessary. The extra effort required to serve several masters would appear to be rather marginal in such circumstances and would certainly not justify the commissioning of

several independent investigators. But measuring physical quantities in a given system is a distinctly different exercise from measuring attitudes of people. If an investigator is embarked on the former, both he and the system are quite indifferent as to the frequency of or even the sequence at which the measurements are carried out (so long as the chosen sequence does not conflict with the need of a particular user to have the measurements taken at a specified time), 'indifferent' in the sense that the results will not be affected by the frequency and sequence chosen. In the case of attitude surveys, however, frequency and sequence may and do have a profound effect on the results. When an employee is asked a question, he does not erase from his memory all the previous questions he has been asked or the answers that he has given in the past, and, either in an attempt to maintain consistency or in order to emphasize certain issues, his reply may be very different from when his mind is not encumbered by such memories. Similarly, there is a limit to which you can persist in probing a human system and a point is soon reached when a higher frequency, which is ostensibly aimed at eliciting more information for the benefit of a wider circle of users, becomes self defeating, leading to increasingly unreliable results and even to undesirable after-effects in the system itself. Furthermore, as the manner in which questions are formulated may reflect the needs of a particular user, it is doubtful whether a single investigator can be so flexible, so detached and so unbiased as to be able to serve faithfully many diverse needs.

We have thus established that "Who wants to know?" is relevant, and we should now consider the question "Why does he (they) want to know?" The obvious answer must be that he wants to determine whether something should be done to change the system. It is argued that management, for whom many of the workers' attitude surveys are carried out, is anxious for the workforce to behave in a way that is compatible with the objectives of the enterprise and which is not detrimental to its smooth running or to certain performance measures. Now, if management accepts the hypothesis that attitude is an important determinant of behaviour, it follows that certain facets of negative attitudes need to be avoided, or changed by appropriate action, hence the supposition that attitude surveys commissioned by management are largely action motivated. I say 'supposition', because in the absence of any intention or desire to act irrespective of the results, there would seem little point in carrying out a survey. Alas, such logic is not always discernible when surveys are embarked upon and all too often one finds that they are conceived with very amorphous objectives in mind, sometimes even as mere delaying tactics, resulting in rather expensive and pointless exercises—pointless but not altogether harmless, since no survey leaves the system totally undisturbed and on occasions it can rouse expectations which, if unfulfilled, can lead to disillusionment and resentment. One cannot, therefore, overstress the need for counsel and for a clearly defined purpose before an attitude survey of employees gets under way.

Having reasonably satisfied ourselves with answers to the questions "Who wants to know and why?" we may now turn to the much more difficult issue of whether the attitude survey can achieve its purpose of revealing how people feel about their work. Whatever the reasons for a person holding certain attitudes (and the reasons are important, since management would naturally want to know whether these attitudes are the outcome of the work environment itself or whether they are shaped by factors outside the managerial competence and control), the problem of eliciting them remains, and the investigator needs to contemplate on the quality of the information that his survey yields and on what assurance he can give that the results are what they purport to be. More specifically, in a given environment and as a result of many influencing factors, a person acquires or develops certain attitudes; the task of the investigator is to determine what these attitudes are, to probe and unearth what is hidden under mounds of chaff, in short to reveal and truly record reality. The question that arises is whether the picture presented by the survey methodology does indeed resemble reality and what possible sources of distortion are likely to be encountered on the way.

In evaluating the outcome of the survey, the investigator needs to consider the following factors that may have a profound effect on the resultant picture:

Vol. 5
1977

(1) Perception;

(2) Articulation;

(3) Identification of explicit, latent or generated attitudes (the observer effect);

(4) The problem of persistent versus transient attitudes;

(5) Interpretation of the measurement concept;

(6) Motivation.

The first factor is an individual's perception ability, or: how does he know what attitudes he has? Since he is an agent in the investigator's probing process, the implicit assumption is that the subject can perceive, with the aid of the questions put to him, what his real attitudes are. Furthermore, he must be able to articulate what he knows (or discovers) about his attitudes. These two attributes of *perception* and *articulation* are essential if the investigator is to regard the subject as the periscope through which to peer in his search. It is because these are attributes that he cannot always rely on, and they certainly vary from one individual to another, that the investigator often has to use his judgement as to how accurate the articulated results are and then resort to making his own inferences in conjunction with his final analysis.

The third factor is the difficulty in separating between *explicit, latent and generated attitudes.* Explicit attitudes are those that the subject is patently aware of, so that after overcoming any initial sense of shyness and reluctance

to respond to the investigator, these explicit attitudes come readily to the fore (assuming the subject is articulate enough to express them, whereas his latent attitudes are revealed only after further questioning and are attitudes which the subject is not aware of prior to the survey). In addition, there could be a set of attitudes which are generated by the manner in which the questionnaire is designed, even by the way in which the questions are asked. That the distinction between the three types of attitudes may be of some importance does not tell the investigator how to separate one from another, and he can never be certain, therefore, that the final picture produced by the survey does not contain certain distortions, many of which may inadvertently be of his own making. Thus, the *observer effect* is an important factor in this field. Even the most innocuous, the most bland and the most gentle investigator is likely to leave some mark on the system under scrutiny, let alone the effect of an insensitive investigator with a strong personality.

In the same way that it is difficult to distinguish between attitudes held initially and those generated by the survey itself, it is often difficult to ascertain the degree to which the attitudes are *persistent* or *transient*. The many factors that shape attitudes are not constant, and variations in mood, as well as recent events (however trivial) in the life of a given subject, may well affect his responses. The transient element in these responses is equivalent to noise generated by atmospheric conditions and which disturbs the quality of transmitted music, but how do you get rid of this noise? The answer appears to lie in repeating the survey, so that the common elements of two (or more) surveys carried out at different times will consist of persistent attitudes, whereas elements that emerge in one survey but not in others may be regarded as a passing phase and hence less significant. The trouble with such a methodology is that—unlike when it is applied to static physical systems—the respondents have memories, that (as stated earlier) every survey leaves a residual effect on the system, that to eradicate these factors it is advisable to allow long intervals between surveys—an impractical solution, since work systems are rarely static over time. It follows, therefore, that determining the transient elements in a respondent's attitudes is often a matter of conjecture and inference.

Another factor in our consideration is the way in which the subject understands the questions put to him, particularly when the answers imply an *interpretation of a measurement concept*. "Taking everything into account, do you feel happy in your job?" How is he to interpret such a question? The investigator can try to elucidate by providing a numerical scale, say 9 indicating a delirious state of happiness to 1 denoting unmitigated misery, and invite the subject to choose a number on the scale; or he may show him a line, the extremes of which correspond to the 9 and 1 states and ask him to indicate a point in between that reflects his state of mind in relation to these extremes; or he may provide a list of possible answers to choose from, such as: deliriously happy, very happy, happy, indifferent, unhappy, miserable. All these methods, whether they are

linked to a numerical/linear scale or whether they are merely ranking of alternatives, are measurement yardsticks, and questions of interpretation are bound to arise. These issues are serious enough, but they are further exacerbated when the question posed to a subject requires estimates of ratios or frequencies. It is well known that questions such as "What proportion of your time do you spend on activity *X*" are difficult to answer and that some people make estimates which are far off the mark; but what about the question "How often do you feel particularly unhappy about your work?" Is the subject to interpret "how often" as a frequency (such as 'five times last week'), or as a percentage of elapsed time? And even if this particular ambiguity can be eliminated (and it can, with further elucidation), are we justified in assuming that the subject understands the proposed or implied yardstick, and furthermore that he is capable of using it?

Another aspect of this problem is that it is much easier for a person to express an attitude to something that he has actually experienced as opposed to something which is rather hypothetical. For example, if you ask an operator, whose work is currently inspected by a quality control engineer, "Would you prefer your work to be inspected by your mate and for you to inspect his?" you require him to rank two alternatives, one of which is currently in operation and with which he is fully familiar, and another of which he has no experience. It is difficult enough to interpret yardsticks and ranking procedures with respect to factual information; when it comes to attitudes in relation to future events, the investigator often finds himself in a realm of pure speculation.

Vol. 5
1977

Finally, we have to consider the *motives* of the respondent. I do not refer to the presence or absence of motivation in his work environment (and there is little doubt that such motivation contributes to the formation of attitudes). My concern here is about his motives in choosing certain answers, thereby deliberately or subconsciously disguising his feelings about his work, in the hope of affecting the results of the survey with some objective in mind. Now, I stated earlier that a survey needs to be conducted for a purpose, that when management commissions a survey (and in most cases it is management that initiates such exercises) it intends to evaluate the results with a view to determining whether to institute changes in the work system. Should we find it surprising, then, that when a subject senses what the results of the survey could be used for, he becomes inclined to produce responses which would lead to favourable changes from his point of view? I think not. It was Bertrand Russell who said that "the two things most universally desired are power and admiration"[1], and it would be idle to pretend that when it comes to attitude surveys all or most respondents assume a placid posture and cease for a while the pursuit of their own objectives.

It is for all these reasons that results from attitude surveys should be con-

[1] See *In Praise of Idleness and Other Essays*, p 34, Unwin Paperbacks, 1976 (first published in 1935).

sidered with great caution. While most conductors of surveys will readily acknowledge the various points made here, many do not seem deterred from embarking on very elaborate and expensive studies and, as a consequence of their understandable commitment to their field work, they tend to advocate too much reliance on the results. It is time that we recognized not only that the attitude survey is a rather simplistic and overrated methodology, but that it can be a hazardous minefield, in which the casualties are rarely the investigators themselves.

SAMUEL EILON
Chief Editor

l. 5
77

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

FEEDBACK

How to Charge for Management Services

IN HIS editorial Professor Eilon [1] puts forward some strong arguments about the difficulties of finding the correct charge to make for management services. While the arguments are persuasive, his conclusions do not inevitably follow from them. My red pencil would hover over the red tape which is implied by his solution—who wants managers to manage the management scientists in management services?

There are, of course, many different types of organisation structure and styles of management, and the organisation of management services should reflect this. In a highly centralised organisation, where the running of subsidiary units and departments is closely controlled and directed by head office, management services may form an important part of the strategic means for exercising that degree of control. In such an organisation head office could reasonably bear the cost of the services it imposes on its operating units. My subsequent remarks have more relevance to the less highly centralised organisation. If the operating units are given any choice in the matter of the use of management services, then should not their choice be tempered by the cost of their decisions?

Before tackling the question of what an appropriate price should be (and the answer could be zero) we must consider the requirements for a pricing scheme. Firstly, we require that our pricing system operates in a way that allocates our resources (the men and machines engaged in management services) most efficiently. Secondly, we require that the costs are recovered fairly. To fit these requirements many different pricing schemes can be used and Eilon mentions some of these: marginal cost, variable cost, average cost, two-tier pricing etc. Out of these, the choice of a blanket zero price, as advocated by Eilon, is not a 'neutral' choice. After all, operating units incur an internal cost, particularly in managerial time, when initiating a project. If the aim is to encourage the use of management services, why not charge a negative price to compensate for these internal costs?

What, then, are the costs of running a project? There are a number of reasons why the average financial cost of a project may exceed the resource value to the operating unit on its own, but not to the organisation as a whole. We can disregard from the discussion those projects where the net benefit is so large that the project will proceed regardless of the pricing system. What we must examine more carefully is the effect at the margin; those projects where the go–no go decision is very sensitive to the price charged. For such projects it is important to ensure that the marginal cost (to the operating unit) should not exceed its marginal benefit at the same time ensuring that the marginal cost (to management services) of running the project should not exceed the marginal benefit to the whole organisation. Such a condition is necessary for all parties to gain from undertaking worthy projects, and to have the incentive to do so. For example, where there are benefits from the project to head office over and above those accruing to the operating unit the project may flounder unless head office bears its share of the costs. Such benefits might arise for any of the following reasons:

1. Head office may identify, more closely than its operating unit, with short-term profit maximising or cost minimising. A small improvement against this objective may be relatively more important to head office than to its operating unit.
2. Long-term benefits may arise; head office is more likely to be prepared to plan and budget for these than the operating unit, whose shorter planning horizons may not adequately account for the long term benefits.

3. Spin-off benefits to other parts of the organisation may arise; organisational learning may occur which will enable other operating units to cope with similar problems more successfully in the future; management services will themselves be better able to deal with such problems.
4. The project may aid head office's primary control function and thus be complementary to other head office activities.

Head office may well accept the case for subsidising management services to the extent that it derives those benefits. A difficult, practical problem arises in trying to assess the amount of benefit. However, that in itself is not sufficient justification for waiving all the charges. Such a policy could have some unwarranted consequences.

Imagine the effects over a period of time of free management services. In the first instance, there would be a tendency for unworthy projects to be started (unworthy in the sense that, if the real costs were known and calculated, they would be more than the total benefits). A second, more insidious, effect would be that all projects, once started, would go on too long. Few projects have a definite stopping point. Whether to continue or stop is an ever-present problem throughout the life of a project. In theory, the operating unit will choose to stop when its expected marginal costs of continuing are greater than its marginal benefit. How will it react when its marginal cost is zero? Like leaving the tap running while brushing one's teeth, it is likely to continue pouring resources down the drain long after it has ceased deriving any real net benefit from them. The third effect is a longer term one. As demand for management services expands beyond the resource optimising point, so there will be pressures to expand the size of the management services department. Once enlarged, the necessity will arise to keep the resources fully employed, and so the artificially stimulated demand will become institutionalised. It will no longer be possible to assess the 'free market' level of demand.

To return to the question of the appropriate charge for management services I have argued that the reason for head office subsidising management services is not so much to encourage its operating units to use management services in general, and the internal services in particular; rather, a subsidy should be applied because head office receives benefits beyond those received directly by the operating unit. Such a subsidy is not a "distortion of the price mechanism", it is a means for head office to pay for the value of services it receives; the 'subsidy' thus restores the effective operation of the price mechanism.

It is not a practical proposition to attempt to calculate the appropriate subsidy on each separate project. It is more realistic to expect head office to be able to assess, in general terms, the proportion of net benefit from management services which contributes to its own (rather than the operating units) objectives, that is, the proportion of management services costs which head office should bear. Charging the remainder of the costs to the operating units will ensure that market demand will reflect the true costs and benefits of projects. If the charges are above this level, demand will be below the resource optimising point and both the operating units and head office will suffer. I have already suggested what the consequences might be if the charges are lowered below this level. How this subsidy is best administered is another matter altogether. It could, for example, be given as a block annual amount, or applied as a percentage to each project. The latter method would act as a greater incentive for management services to stimulate its own demand. If management services were, additionally, a cost and profit centre in its own right, the necessary conditions would have been created for the demand for management services to be self-regulating. Many of the supposed advantages of centrally directed management services could thus be present in a decentralised system. In particular, the responsibility for initiating projects, when some degree of prior technical knowledge is needed, could be confidently placed in the lap of management services itself.

I have sought to argue the case for maintaining market forces as a means of regulating management services. These arguments will carry varying weights in different kinds of organisations. But that said, the appropriate conditions can be created for management services to be self-regulating. The alternative is an administrative system for organising, controlling, motivating and evaluating management service activities. It is hard to see how the cost of running such an adminis-

Vol. 5
1977

trative system can compensate for any imperfections which might arise in the market mechanism. Let the administrator prove his worth before control is relinquished to him, and let management services demonstrate that a genuine market demand exists.

RAY TOMKINS
Department of Management Science
Imperial College
Exhibition Road
London SW7 2BX

REFERENCE

1. EILON S (1976) How to charge for management services. *Omega* 4(6), 623–628.

1. 5
77

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

Employee Attitude Surveys: Their Interpretation and Use[1]

DG HARPER

Imperial College, London

T KYNASTON REEVES

Anglian Regional Management Centre,
North East London Polytechnic

(*Received June* 1976; *in revised form November* 1976)

This paper evaluates the uses of attitude surveys as a guide to management decision-making and identifies some of the problems of carrying out such studies and implementing the results. The concepts of attitude, value, personality, motivation, morale and job satisfaction are reviewed and their relevance discussed. A model of the determinants of attitudes is developed and used as a basis for a consideration of the problems of effecting change and increasing job satisfaction. Finally, the types of action that might be recommended on the basis of attitude survey data are considered and the roles of management and researchers in the interpretation and implementation of conclusions discussed.

INTRODUCTION

STUDIES of human relations within large organisations usually involve some kind of systematic attitude survey. The use of attitude surveys raises a range of problems associated not only with measurement of attitudes but also with the psychological processes underlying the attitudes being measured. In this paper we consider some of these problems and their implications. Resolution of these issues is not merely of theoretical interest but has a direct bearing on how the results of attitude surveys are used as a basis for action.

Although there is a considerable literature on attitude surveys, in general very little has been written on the particular problems of carrying out attitude surveys in organisations. This article aims to distil some of our experience as practitioners in the field which might be useful either to those embarking on this kind of organisational research or to those in organisations concerned with making use of the results.

[1] The authors wish to express their thanks to the National Coal Board, Black & Decker Ltd., Cadbury Schweppes Ltd. and British Airways for their cooperation in the publication of the survey results presented in this article.

We wish to argue that attitude surveys can be a valuable guide to management action provided they are not used simplistically and the user is aware of certain cautions in interpreting and implementing their findings.

SOME CONCEPTUAL CONSIDERATIONS

Attitude surveys are basically a systematic way of finding out what people are thinking and feeling. In our everyday interactions with other people we are constantly making assessments of this kind. We need only turn to attitude surveys when we are dealing with large groups of people in which it is not possible to get reliable information on their attitudes or feelings simply from one's direct personal contacts.

In the course of any discussion of attitudes one finds a number of other terms used: personality, motivation, morale, values and opinions. It is appropriate to consider some of these terms briefly to clarify what we mean when we are using them. A further term we shall need to clarify is 'job satisfaction'.

Vol. 5
1977

Attitudes and Values

An attitude may be described as the individual's psychological relationship with an object or an idea. Katz [2] proposes the following definition:

> An attitude is the predisposition of the individual to evaluate some symbol or object or aspect of his world in a favourable or unfavourable manner.

This definition carries the implication that attitudes are concerned with evaluations and will therefore be indicators of the individual's values. These evaluations will have both an emotional and cognitive component. In other words the individual will have positive or negative feelings towards an object or idea and also certain views or beliefs about it. For example, two people may both have negative attitudes towards incentive payment systems, but their reasons (the cognitive element) for holding that attitude may be very different: one of them may dislike them because the weaker or less skilful worker is penalised, whilst the other may dislike them because they lead to a decrease in the quality of work. It follows that it is important to know the cognitive content as well as the direction of the attitude: this will be relevant to the process of attitude change since the information on which the attitude is based may not be well-founded. These cognitive and emotional components of attitudes will be reflected in the formulation of the attitude questionnaire, in that the questions will usually ask how a person feels about a particular thing and then elicit his reasons for feeling that way. An assumption of attitude theory is that, all things being equal, the holding of a certain attitude will lead to behaviour that accords with it. In real life of course all things are not equal and many factors intervene to prevent

behaviour being determined exclusively by a particular attitude a person holds.

The term attitude survey and opinion survey are often used interchangeably. Strictly speaking, this is incorrect: an opinion is concerned with beliefs about a matter of fact (e.g. "Who do you *think* will win the next election?"), whilst an attitude is concerned with a statement of value (e.g. "Who do you *want* to win the next election?"). In practice, both attitude and opinion questions would both be found in an attitude survey.

Personality traits

Personality traits are distinguished from attitudes in that they are not concerned with the individual's relationship to particular objects. By personality we understand the individual's typical way of behaving. For example, a person who has an aggressive personality will tend to react aggressively in a wide range of situations. If we know an individual has an aggressive disposition, this knowledge by itself does not allow us to predict what attitudes he holds; it does tell us, however, that whatever his attitudes are he will tend to express them more aggressively than the average person. Thus when we talk about personality we are concerned with differences between individuals in typical ways of responding. Within the context of a personnel selection programme, for example, it could be very important to measure these differences. But although personality may influence attitudes, it is not usually of relevance within the framework of an employee attitude survey where the action that is likely to be taken will not take account of this kind of difference between individuals. For the purpose of policy formulation—as opposed to day-to-day dealings with people—it is not possible, or perhaps desirable, to make distinctions on the basis of personality. It is not, for example, realistic to have one set of policies for extroverts and another for introverts.

Motivation

An important aspect of personality that is relevant to attitude surveys, however, is motivation. For example, important motives in work situations are the needs for affiliation, achievement, power, and recognition. People will vary in the relative levels they possess of each of these, and these variations may be related to job performance; for example, there is evidence that successful managers tend to have higher than average needs for achievement and power. In an attitude survey it is likely to be important to establish the dominant motives or needs of the group or groups being studied and to examine how far these needs are being met by the work situation. This can be a valuable indicator of job satisfaction.

The emotional component of attitudes referred to earlier is in fact the attachment of these motivational drives to specific objects. It is with these objects and the social processes through which drives become attached to them that we are primarily concerned in attitude surveys. The critical point that we wish to

make here is that while the level of motivation which forms part of a person's personality may remain stable, the objects towards which it is directed may change. When a motivational drive becomes attached to an object, it is then experienced as a need for that object.

A discussion of needs raises what we consider to be a critical issue, namely how far it is possible to use attitude surveys as a means of identifying the strength of needs which may be latent, as contrasted with needs which are manifest and to a greater or lesser extent are being currently satisfied. In effect, we must envisage the possibility that needs exist which the individual is either unaware of, or which he does not conceive of being satisfied within the work context. For example, how can someone who has never participated in decision making, know that he will derive satisfaction from such participation? In an important sense, a major problem in both consumer research and employee research is the identification of these latent needs. However, it may be the case that the identification of them must be guided by a theory of human motivation rather than by research.

Morale

Morale is a term frequently—and loosely—used in industry to refer to the state of employees' attitudes and motivation. More precisely morale might be defined as the level of motivation with regard to a set of goals [6]. Although the possession of a high or low level of motivation is essentially an individual phenomenon, the term morale is perhaps more often used to describe the general level of motivation of a social group. Thus a state of high morale refers to a situation where there is, collectively, a high level of motivation to achieve the group's goals, while a low level of morale describes a situation where there is, collectively, a depressed level of motivation. The members of the group are usually aware of their state of morale subjectively, but in common with other motivational states it is difficult to measure.

The existence of a state of high morale among a group of employees does not, however, necessarily mean they have a high level of commitment to achieving organisational goals. It is of course management's aim to harness motivation in this way, but a state of high morale can equally well exist in relation to goals that run counter to those of the organisation, as for example in a strike.

High levels of absenteeism, labour turnover, accidents and disputes, in that they are manifestations of avoidance of work and withdrawal of commitment or interest, are often interpreted as behavioural indicators of low morale. Attitudinal indicators of low morale are usually taken to be synonymous with low job satisfaction. One would normally expect attitudinal manifestations to be accompanied by behavioural ones [4], though circumstances may prevent people from expressing their dissatisfaction in overt behaviour, for example in an economic situation where jobs are scarce.

Vol. 5
1977

Job satisfaction

The final concept that we wish to discuss is that of job satisfaction. A great deal has been written and a considerable amount of research carried out on job satisfaction. A useful review article is that of Locke [3] who notes that over 3000 articles on the subject have been published in the last two decades. Locke has defined job satisfaction as

> A pleasurable or positive emotional state resulting from the appraisal of one's job or job experiences.

It is a difficult concept both to analyse and measure. Attitude surveys have consistently identified remarkably high proportions of people who claim to be satisfied or very satisfied with their jobs; in fact between 70 and 90 per cent can be predicted to feel this way regardless of the kind of work being done, and often, also, regardless of the fact that grievances about specific aspects of the work situation may be quite strongly held. Some not untypical examples of a measure of job satisfaction and the distribution of responses obtained from employees in very different jobs and industries are shown in Table 1, which uses data from surveys conducted by the present authors.

It will be seen from this table that, while there are minor variations in job satisfaction between the organisations, the general pattern of responses is similar despite the fact that the respondents were working in very different industries, and also that a high proportion had expressed elsewhere in the questionnaires very specific grievances concerning some aspect of their work situation. Thus, differences in attitude that might have been expected were not reflected in this measure of job satisfaction.

The problem of measuring specific sources of job satisfaction, as compared with the global measure in Table 1, has also received a considerable amount of attention in the literature [3]. Psychologically it seems important to obtain two types of information. First, what aspects of the job are salient or important to the individual, and secondly, how he evaluates his present job in terms of those same aspects. Taken in combination the responses to these two questions will permit an identification of those features of the job which fall short of expectations, and are important to the individual. Clearly, if the job is deficient in some aspect, e.g. overtime opportunities, but this is of little or no importance to the individual then it is unlikely to contribute significantly to job dissatisfaction. Figure 1 shows an example of the relationship between salience and evaluation by a group of industrial workers for 24 job attributes. The factors of the job in the top right hand corner of the diagram are likely to be major sources of dissatisfaction, whilst those in the bottom left hand corner come up to, or even exceed expectations. A diagram of this kind permits an easy identification of specific sources of dissatisfaction.

Certainly, from the point of view of management action, it is often much more useful to measure a person's attitudes to particular features of his job,

TABLE 1. EXAMPLES OF RESPONSES TO A MEASURE OF JOB SATISFACTION[1]

Responses to the question:

"Can I ask you now what you think about your job as a whole? Taking everything into account, the pay, the people you work with, the supervision and the work you do, how do you feel about your present job?"

	British airways ground staff	British airways cabin crew	Factory workers Cadbury Schweppes	Factory workers Black & Decker	National Coal Board Mineworkers
I like it very much	33%	42%	37%	35%	47%
On the whole I like it	56%	55%	52%	54%	41%
Indifferent/No response	6%	1%	1%	0%	0%
I'm not really keen on it	4%	1%	8%	9%	10%
I dislike it	0%	1%	1%	1%	2%
I hate it	0%	0%	1%	1%	0%
Numbers	314	189	586	405	1024

[1] Data taken from studies carried out by the authors.

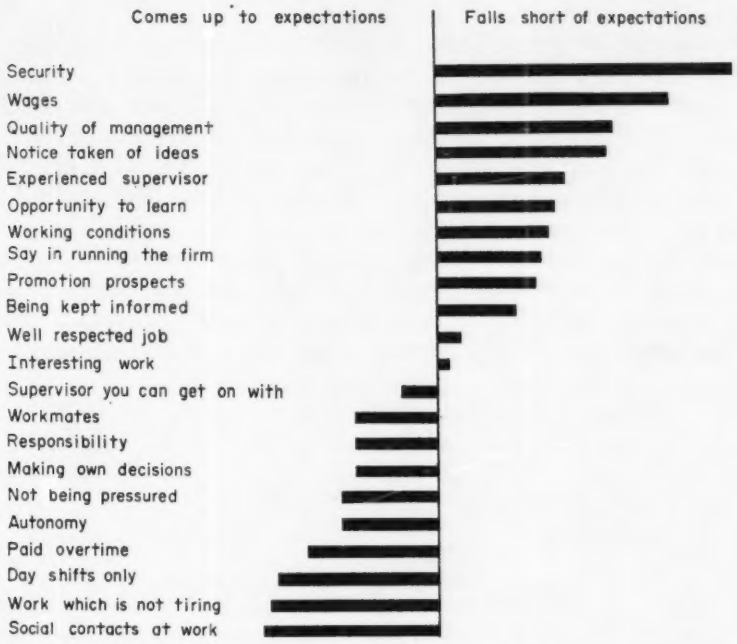


FIG. 1. *Sources of job satisfaction, and dissatisfaction amongst a group of industrial workers (data taken from a study carried out by one of the authors).*

for example, to pay, to security, to interest of the work, to the supervision, to promptness in dealing with grievances, etc. and then to take action to put right those features of the work environment that are causing negative attitudes in any of these respects. A useful discussion of different approaches to the design of composite measures of job satisfaction is to be found in Quinn and Mangione [5]. However, even if specific sources of dissatisfaction can be identified in this way, a means is still required of combining them into an overall measure of dissatisfaction, so that individuals or groups of individuals can be compared, or indeed the same groups compared over time.

THE INTERPRETATION OF EMPLOYEE ATTITUDE SURVEYS

To make practical use of employee attitude surveys two key things have to be discovered. The first is what people's attitudes actually are, and the second what the determinants of these attitudes are. An understanding of these determinants will prove a guide to the action that can be taken to eliminate the causes of negative attitudes.

Determinants of attitudes

In an employee attitude survey we usually wish to explain those aspects of a person's attitudes that relate to his job: namely his attitudes to the various features of his job and his attitude to his job as a whole, that is to say his job satisfaction. Possible determinants of these attitudes are depicted in Figs. 2 and 3. Figure 2 presents a broad overview of these determinants, while Fig. 3 shows in greater detail how a person's more general attitudes to life and to work can bear on his attitude to a particular job—the inner dotted box in the diagrams.

In addition to measuring these 'job attitudes' we also need information which will enable us to explain why these attitudes are positive or negative as the case may be. We do this, in most cases, by reference to features of a person's working environment—his working conditions, his pay, the kind of supervision he gets

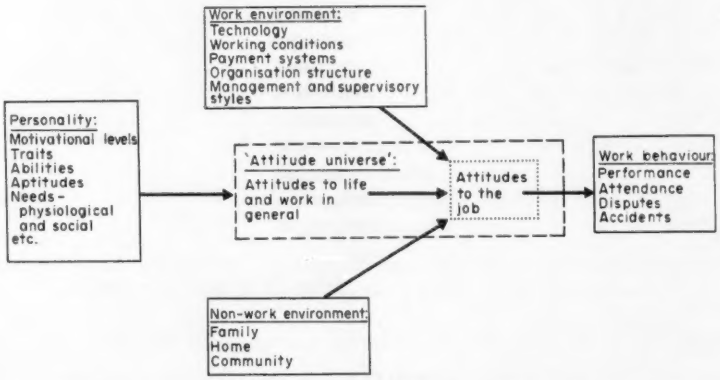


FIG. 2. *Determinants of attitudes to work: non-attitudinal factors.*

and so on [3, pp. 1319–1328], and sometimes, though in practice this seems to be less often done, by reference to his non-work environment—the kind of community he comes from, the number of children he has, the commitments he has undertaken, etc. Indeed it seems likely that these external factors may be growing in importance as determinants of attitudes to work [1]. To the extent that this is so, management's ability to change attitudes by changing features of the work environment that they control is thereby reduced.

As well as being influenced by the various tangible factors indicated in Fig. 2, a person's attitudes to his job will also be affected by other more general attitudes he may hold (Fig. 3). Thus, a person's attitude to a particular piecework system could be determined by the fact that it happens to be a particularly unfair system. It could be, though, that the person has an attitude to piecework systems in general, for example he disapproves of making wages dependent on output, in which case his attitude to any particular piecework system is likely to be negative no matter how equitable the particular system happens to be.

Vol. 5
1977

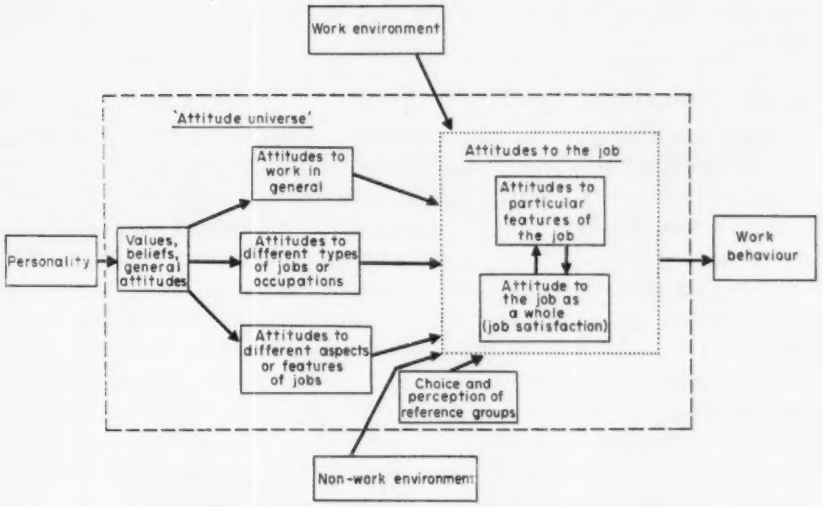

Work environment
'Attitude universe'
Attitudes to work in general
Attitudes to the job
Attitudes to particular features of the job
Personality
Values, beliefs, general attitudes
Attitudes to different types of jobs or occupations
Work behaviour
Attitude to the job as a whole (job satisfaction)
Attitudes to different aspects or features of jobs
Choice and perception of reference groups
Non-work environment

FIG. 3. *Determinants of attitudes to work: attitudinal factors.*

Another very important determinant of attitudes which needs to be mentioned is a person's reference groups. These are those groups of people he identifies with and uses as a basis of comparison for evaluating his own position. Thus, a critical determinant of a person's attitude to his pay, for example, could be whether he compares his money with that obtained by other people doing similar work inside his company or industry, or whether he compares his earnings with those of outsiders.

Referring again to Figs. 2 and 3, the various determinants of job attitudes just discussed have been positioned either inside or outside the outer dashed rectangle. This dashed rectangle contains what might be described as a person's 'attitude universe'. From the point of view of any management action to try and bring about an improvement in attitudes, it is readily apparent from the diagram that the determinants of attitudes most within management's control are the tangible aspects of work environment, outside the person's 'attitude universe'. It is here that management has most scope for action to eliminate sources of negative attitudes and cultivate sources of positive attitudes. The implication of this is that the researcher if he is to be effective should devote a considerable amount of effort to understanding the organisational and technical context within which the population being studied works.

However, although these organisational factors are important determinants of attitudes to work, it does not mean that changes in job satisfaction cannot be brought about by trying to change a person's more general attitudes or his beliefs or values or his motivation. In fact one of the objects of training may be to encourage people to examine and reappraise the bases of their attitudes and

prejudices, and to examine how far they are consistent with their work objectives.

Conscious and unconscious attitudes

One anxiety often expressed about employee surveys is that the answers given may not reflect people's 'true' feelings: either because they are inhibited about telling the truth or because they are not motivated to co-operate fully. In our experience these fears are usually unfounded, for very definite reasons. Work, even if not enjoyed, is normally a matter of central concern to people, and is something about which people have very conscious feelings. Telling others how they feel about their work is an activity which most people seem to enjoy, and provided a sufficient level of trust has been established by the researchers, employees at all levels in an organisation will normally be interested in and co-operate with a survey. Indeed, so much so in our experience, that they will be at considerable pains to express their true feelings, despite occasional shortcomings in the way a question has been phrased. This general consciousness of feelings about work coupled with an interest in the survey and its outcome means that one can normally attribute validity to the responses to survey questions.

On the other hand, not all aspects of a person's feelings about his work are necessarily consciously held, and, more particularly, people are not always aware of the reasons for feeling the way they do. This means that the researcher must be prepared not to take all his results at their face value, a point which has already been illustrated with regard to the measurement of job satisfaction.

This may be further illustrated by reference to our survey carried out in 1971 of shopfloor workers in Cadbury-Schweppes Bournville Factory, a large confectionery factory employing at the time some 6000 people. Although much of the work in the factory was repetitive and people had to gear their movements to the pace of a machine or moving belt, only a very small minority of the sample expressed overt dissatisfaction with this kind of work. Indeed, as Table 1 has shown, nearly 90% expressed overall satisfaction with their job. However, when people were compared on a composite measure of job satisfaction (a weighted combination of scores on a number of attitude scales including such items as: 'How often do you feel you would rather stay away from the job than come in?' 'How often do you feel particularly dissatisfied or frustrated by your work?'), it was possible to show clear differences in response between people doing different kinds of work, with those doing 'paced' work tending to have very much more unfavourable attitudes to their job than those doing unpaced work. Thus, although no *overt* dissatisfaction with paced work was expressed in the survey, a relatively high level of dissatisfaction could be demonstrated by statistical analysis. Clearly a job enrichment programme designed to reduce the incidence of this kind of repetitive paced work could have been expected to improve job satisfaction.

This well illustrates a problem of interpreting what people say: how can

people realistically assess how far they want something that they have never experienced and possibly do not clearly understand? The opportunity to perform fulfilling work is outside the expectation of most ordinary workpeople, and was therefore unlikely to figure highly among the aspirations of the Cadbury workforce. Indeed, this is an area where management perhaps needs to take a lead in shaping opinion rather than following it. Providing people with the experience of 'enriched' fulfilling work could create its own demand.

Thus, analyses of this kind can reveal attitudes which are either subconsciously held or not articulated clearly, but which may nevertheless influence the individual's adjustment to work.

TAKING ACTION ON SURVEY RESULTS

A major rationale for carrying out employee attitude surveys is that if we know what people are thinking and feeling then we can deal more effectively with their present needs, anticipate their future needs and develop a more satisfied or more involved group of people within an industrial organisation. The objectives in doing this may be either instrumental or altruistic. An instrumental approach entails the view that increasing people's job satisfaction or happiness at work will further primary organisational goals such as improving productivity or reducing labour turnover. The altruistic approach, on the other hand, regards increasing job satisfaction as an appropriate organisational goal in itself, irrespective of whether or not there is any demonstrable relationship between job satisfaction and more effective work behaviour.

The recommendations which might emerge from an attitude survey are of two broad types: first, those that lead to the immediate removal of particular sources of grievance, and second, those that integrate a number of the research findings into some long term programme of change. In drawing out recommendations from attitude survey results there are a number of important issues.

One, which has already been discussed in the context of the Cadbury-Schweppes survey is that people's attitudes to their work may not always be consciously held or overtly expressed. This means that overtly expressed views—in that particular case criticisms of the way work was organised and allocated by supervisors—may take precedence as a basis for action over other sources of dissatisfaction which may be established by analysis of the data but which are not overtly expressed. Much will depend here on the skill of the researcher in presenting the results.

A further issue concerns the proportion that should hold a certain attitude for action to be justified. Should action be taken if only a few people express a grievance, should it be taken if 50 per cent express it, or should it be taken only if the majority of respondents express it? The absolute numbers involved will of course be relevant here. However, the attitude survey itself cannot provide

an answer to this kind of question and it becomes a matter of managerial judgement. What management decide to do will be influenced though by the way the researchers interpret the results, and it is at this stage that the attitude survey becomes something more than a means of upward communication of employees' views and opinions.

Assessing the significance of changing trends in morale

These points can usefully be illustrated by reference to a series of studies of mineworkers in the British Coal Industry. Figure 4 shows the responses of samples of mineworkers to the same attitude question asked at intervals over a

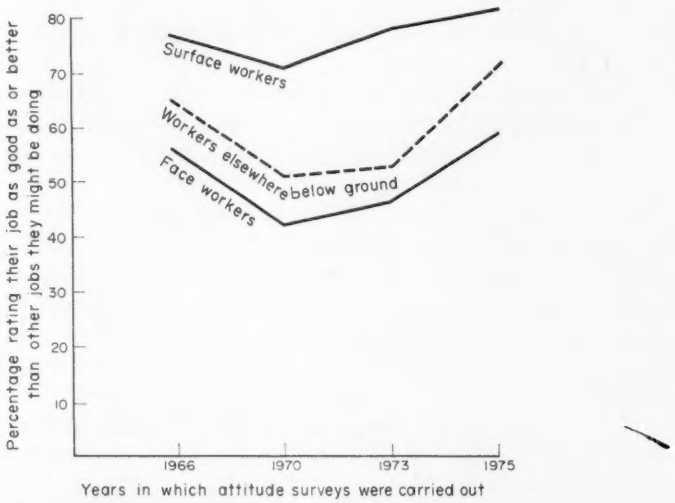


FIG. 4. *Responses to the question: "Looking at it overall, how would you say your job compares with other jobs you might be doing outside mining?" (Data taken from 4 studies carried out in the National Coal Board by one of the Authors.)*

period of 10 years. The question is designed to provide a measure of the mineworker's overall evaluation of his job; in order to do this it asks for a comparison of his job with others that he might be doing in other industries. The question is of course a very general one and many factors could influence the evaluation made: what is important in the present context are the trends up or down in each of the three groups of workers.

Two main points emerge from this graph. First, between 1966 and 1970 there was a substantial increase in the proportion of mineworkers rating their job negatively. Secondly, the deterioration in attitude was greatest amongst the more influential groups of mineworkers: amongst underground workers this change was 15 per cent compared with only 6 per cent amongst surface workers.

Vol. 5
1977

A further point to be made here is that at each of the points in time morale was lowest amongst face workers—the most influential group of all.

There was some evidence, therefore, in 1970 that industrial problems were likely to occur unless the downward trend was checked by appropriate managerial action. In fact, in 1971 a national strike occurred in the industry and it is reasonable to interpret the trend in Fig. 4 as the early warning signs of the unrest that led up to the strike. The subsequent pay settlement reversed the downward trend but it was only by 1975 that evaluations of the job were restored and increased above the relatively favourable level of 1966. This trend presumably reflects the continuous improvement over this period in pay and conditions in the industry. The evidence suggests that a confrontation such as a strike may have long term effects on morale that take a long period of time to remove.

There are, however, two problems in interpreting and acting on this kind of data.

First, the change observed between 1966 and 1970 was 15 per cent in the most extreme group: is this change sufficiently large to indicate action to reverse the trend? The first step that needs to be taken is to establish the statistical significance of the change; if it proves that it could not have occurred by chance, then the relationship of the survey data to other factors, e.g. labour turnover figures, industrial relations disputes, management–union relationships —should be examined. Basically, if all of the available data are consistent with the morale data, then the need for corrective action is strongly indicated. The second step is to analyse the relationship of the trend to the more detailed attitude data: this may enable the factors most strongly associated with the trend to be identified. These factors may vary from one group of employees to another: for example, whilst all employees may be dissatisfied with pay levels, the situation may be aggravated in one group by boring work and in another by lack of promotion prospects.

The second problem stems from the fact that attitude survey information is only one factor to be taken into account by the chief executive or other individuals responsible for the relevant decisions. Perhaps the most important of these other factors will be ongoing bargaining or negotiating between management and the unions: whilst the attitude survey will provide important information on the strength of feeling amongst the rank and file on relevant issues, the complexity of the bargaining situation may prevent it from being acted on in any simplistic way.

Management's role in interpreting survey findings

Developing this general issue of the translation of survey data into action, a course of action can only be determined when the data are considered in relationship to a managerial or social philosophy. This is especially relevant when

considering long term personnel programmes. Perhaps this point will become clear if we illustrate it with an example from a survey of Black & Decker employees carried out by one of the authors. The survey identified considerable discontent with the perceived unfairness of the way the payment system was administered. Two lines of action were open to management: either to change the payment system to fit in with the workers' requirements, or to continue with the system in spite of the grievances. In this case, because the system was consonant with the broader managerial philosophy of the firm and was felt to lead to high productivity, management did not feel able to follow majority opinion. However, the attitude survey results made them far more aware of the negative consequences of their payment practices. The issue, central to the management of organisations, that emerges here is that of striking a balance between the needs of the individual and the overall organisational goals and policies. Attitude surveys serve the function of making this kind of problem explicit and thus increasing the likelihood of it being dealt with satisfactorily.

A further important point is that any set of actions that are recommended will bring expected benefits, but are also likely to have costs. There are limitations, however, to the extent to which it is possible to evaluate the recommendations in cost–benefit terms. For example, in a recent survey of British Airways staff a need was identified for more face-to-face communications from management. In order to achieve this contact employees might attend on some regular basis seminars in which management could communicate information about problems, policies, and proposed changes, and employees would have an opportunity to express their views and engage in discussion with management. A series of seminars of this kind could be easily costed in terms of employees' and management's time and the cost of facilities; the benefits, however, in terms of improved communications and consequent gains in terms of involvement in and commitment to the job cannot be costed in financial terms as in cost–benefit analysis of, say, a new technical system. In effect, therefore, whilst the cost–benefit analysis will show whether the cost of the proposed recommendation is totally unacceptable the decision will be largely influenced by management's belief in the importance of this kind of action and in the expected benefits.

A final point is that the possible lines of action that might be recommended on the basis of the survey should be given careful consideration at the very beginning since it is unwise to raise expectations with regard to a particular issue if it is one which for any reason management may be unable to do anything about. This leads to the more general point that management must always embark on this kind of activity fully committed in principle to some kind of action and ensure that after the survey is concluded some kind of action is in fact taken: unless this happens the failure to meet the expectations raised by the survey may leave the company in a worse position from the point of view of employee attitudes than if it had not been carried out at all.

Vol. 5
1977

REFERENCES

1. GOLDTHORPE JH, LOCKWOOD D, BECHHOFER F, and PLATT J (1969) *The Affluent Worker in the Class Structure*. Cambridge University Press.
2. KATZ D (1960) The functional approach to the study of attitudes. *Pub. Opinion Q.* **24**, 168.
3. LOCKE EA (1976) The nature and causes of job satisfaction, in *Handbook of Industrial and Organisational Psychology*. (Ed. DUNNETTE, MD), Chapter 30. Rand McNally, Chicago.
4. PORTER LW & STEERS RM (1973) Employee turnover and absenteeism. *Psychological Bull.* **80**(2), 151–176.
5. QUINN PR & MANGIONE TW (1973) Evaluating weighted models of measuring job satisfaction: a Cinderella story. *Org. Behav. Human Perf.* **10**, 1–23.
6. VITELES MS (1953) *Motivation and Morale in Industry*, p. 284. Norton, New York.

ADDRESS FOR CORRESPONDENCE: *Mr DG Harper, The Industrial Sociology Unit, Imperial College, 52/53 Prince's Gate, London SW7 2PG.*

l. 5
77

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

Reflections on Operational Research. A Report from the Euro II Congress[1]

MALCOLM DANDO

The University of Sussex

COLIN EDEN

The University of Bath

(*Received February* 1977)

The EURO II Congress at Stockholm contained a set of Invited Papers. The set included a group that appeared to be attempting a major review of the role of Operational Research (OR). We try here to give a short summary and a representative selection of extracts from these papers. This is done partly because the papers will not be easily available to all OR workers who are interested, but mainly because of the emphasis they place on the need for truly interdisciplinary OR which recognizes the complex and potentially partisan nature of what we do. Related work in the UK is noted in order to emphasise the beginning of a widespread concern over such issues. The remainder of the Congress demonstrated a distinctly different view of OR, which was characterized by a disregard for the nature of human behaviour and of power in organizations and society. We suggest that the reflections made by the Invited Speakers demonstrate again the need for a fundamental reorientation of OR.

INTRODUCTION

THIS is an attempt by two participants from the United Kingdom to report on a part of the Second European Congress on Operations Research. The qualification—that this is a partial report—is important because the Congress seemed to us to proceed in two parallel, but quite distinct parts. One part is available in the Abstracts and Proceedings of the Congress. The other part of the Congress consisted of a set of Invited Papers from Scandinavian participants. The nature of the distinction between these two parts seems to us to be important, and we report below on our analysis of this difference and argue that the results of the analysis have major implications for the practice of OR.

In general the Invited Papers were probably very close to the "Research in OR" that Mitchell [18] called for. In outline, the following review attempts to highlight the complexity of the problems OR workers are now being asked to

[1] Stockholm, Sweden. 29 November–1 December 1976.

study, the role that such studies play in their organizations or society's activities, the difficulty of assessing very complex situations, and, perhaps most fundamentally, the varying assessments that will be made of the 'same' situation by participants having differing interests.

REVIEW OF INVITED PAPERS

The Invited Papers are listed in Table 1 and will be dealt with in the order shown as that is how they were arranged at the Congress. We shall deliberately

TABLE 1. INVITED PAPERS AT EURO II[2]

1. Sven-Eric Andersson, President of the Swedish Operations Research Association.
   Past Results and Future Prospects. Welcome Address

   Lieutenant General Lennart Ljung, Chief of the Swedish Defence Staff.
   Generals Need Help to Think—OR to Structure Defence in an Uncertain and Changing World

   Arne S Lundberg, Former President of LKAB, Sweden.
   Operations Research and the Say of a President

   Gunnar L Johansson, Senior Vice President of The Volvo Group, Sweden.
   Operations Research and Socio-Technical Development

   Professor Arne Jensen, The Technical University of Denmark.
   Future Prospects of OR

2. C-A Staël von Holstein, Senior Decision Analyst at the Stanford Research Institute, Stockholm.
   Quantification of Judgement and Preferences in Decision Analysis

   B Brehmer, Professor of Psychology at the University Uneå.
   Human Judgement and Decision Processes

3. K Nygaard, Director of Research at the Norwegian Computing Center, Oslo.
   Cooperation with the Trade Unions—A Challenge to Operations Research Workers

   V Epstein, Training Director of the Swedish Employers' Confederation.
   Industrial Democracy—Impact on Management Techniques

   Å Sandberg, Research Assistant at the DEMOS Project and the Dept. of Sociology, University of Uppsala.
   The Role of Rational Analysis in Decision-Making—Harmony or Conflict Perspective?

[2]The Table only includes those relevant to the present argument. Other speakers were Åberg, Björkman, Jansson and Näslund and in some instances those may also have been relevant, but we missed their presentations due to parallel sessions and have not been able to read their papers. Available papers were published at EURO II, as edited by B Björklund, Royal Institute of Technology, Stockholm, to whom we are indebted for his kind assistance.

Vol. 5
1977

use quotations in order to convey major points in the words of the original author. It should be understood, however, that each author was using examples from his own area of expertise to draw some general conclusions about OR. For present purposes we have merely attempted to sketch in the context and have concentrated on the general conclusions.

Sven-Eric Andersson the President of the Swedish OR Association set the scene;

> "At present, much effort is devoted to the development of budgeting and strategic planning processes by support of OR within public sector as well as in industrial sectors. Many problems partly new to OR then face us. Often, the mathematical problems themselves provide formidable tasks. Always, the successful integration of OR achievements into existing organisations and their procedures require extensive sales and implementation efforts."

This theme was taken up immediately by General Ljung, the Chief of the Swedish Defence Staff. After laying out the reasons why his organisation was interested in long term planning and how this was organised in Sweden's Defence Establishment he turned to OR and its role. Almost his first remark on this subject was;

> "It should be mentioned that the term operations research in the Swedish Defence today has a broader meaning than what you probably have. We include systems analysis and, to some extent, also economics and policy science in our conception of operations research."

General Ljung outlined the organisation of OR analysts in the defence system, and the earlier point about the scope of OR was then reinforced;

> "Since 1958 when the organisation was established, there has been a shift in the problems studied by the analysts, from questions of procurement to long-range policy problems.
> In the same time the fields of science from which methodology is acquired have been extended from mathematics and natural sciences into encompassing also economics, political science and organisational and behavioural sciences."

An example of a military planning problem was then given with the conclusion that;

> ". . . I have tried to illustrate the fact that our OR experts do a lot more than building models. They take part in the formulation of problems and goals. They participate in generating alternatives and in choosing criteria. The example also touches on structuring problems and discussions on 'softer' methods to be applied when the questions become more politically tinged."

Finally General Ljung presented his views of OR (from the perspective of a decision-maker);

> "It must be emphasised that studies do not provide us with the correct answer. The studies are there to increase the knowledge of the problem considered and by that to improve the decision to be taken. . . .

The results of studies should not be presented as true answers but as conclusions drawn from causal connections and assumptions."
". . . you bring your education and values into our military organisation. Since your approaches and lines of action differ from ours, you may influence our attitudes and develop our ways of thinking: Generals do need help to think!"

Bishop [2], Bowen [4] and Eden and Harris [12] have, for example, come to very similar conclusions with regard to OR for complex decision-problems. A short presentation followed from AS Lundberg, former President of LKAB Sweden. Mr Lundberg began by identifying the uncertain world of top management;

". . . No wonder probability calculation is relevant at the top, because there you do not live in a world of exactitude. The only constant probability is that you go bankrupt, if you do not act."

The relationship between this businessman's problems and that of the General seemed rather clear—as did the kind of help that they thought was useful. Mr Lundberg's advice to OR men was also direct;

"So my advice to the operations analysts when you present the results of your toil to so-called decision-makers—who usually decide what you tell them—is you should use most of the time to discuss the basic assumptions to see that as few factors as possible have been forgotten and that the decision-makers agree on the questionable input."

Vol. 5
1977

From this discussion of the role of analysis in high-level policy making our attention began to be directed to examples of the impact of OR-aided decisions on the affected systems. This issue was set out by GL Johansson, Senior Vice-President of AB Volvo. After an outline of the organisation of the Volvo Group of Companies and of the OR group within it, Mr Johansson stated his theme;

"In my address, however, I should like to limit myself to the applications of operations analysis within production, particularly the way in which operations analysis methods have played their part in socio-technical development. . . ."

He also set out his general philosophy and its implications for OR at this early stage;

". . . More and more specialists are providing basic material for decisions. Furthermore, viewpoints from the people who are to carry out the work are being taken into account at an early stage so that working conditions are not determined too much by the purely technical or merely financial evaluation. . . . Naturally enough, decision and operational analysis has a very definite part to play here particularly when the decisions to be made have great financial consequences. There is, however, reason to underline the value of experience and knowledge as well as the significant development which is based on successive and continuous changes and improvements, in other words the know-how of the Company personnel and all those who supply equipment and systems to our facilities. This natural development has the advantage that it can be understood by many and that many people can make their contribution. . . ."

Mr Johansson gave some examples of OR operations in the Volvo Group. The largest amount of time was spent on the well-known Kalmar Plant example. Before proceeding to his other examples the message for OR was spelled out;

> "... Under no conditions may systems be made up and analysed within a very limited group of specialists and then all the new material installed in the form of a finished package. On the contrary, the creation and application of a project requires a maturing and training process in which ideas, system solutions and details should be discussed and criticised. The approval of the Company concerning a project provides the intentions and outlines but leaves many questions to be answered...."

After other examples he concluded specifically;

> "By presenting these examples from our activities on the production side, I have tried to indicate a few things. The decision situations are tending to become more complicated the whole time, there are more factors to take into account and more people are taking part in the decision process. Good basic material for decisions can eliminate many unnecessary discussions and misunderstandings but it is important to obtain confidence in advanced methods and it is also important for the results to be given a very concrete form which is easy to understand....
>
> We have also found it important to use the analysis method in a constructive way. A change in a socio-technical system must be based on organisational co-ordination and flexibility in the work organisation as well as on new methods and technical solutions...."

Thus ended the initial phase of the reflections—each by someone who has to respond to the work of Operational Researchers.

The second phase, of presentations by people within the profession, was started by A Jensen who presented the final paper of the Plenary Session. Whilst we had some difficulty in understanding parts of this paper, there was a great deal of content that was very clear, and easily related to what had gone before. Professor Jensen opened with the central point;

> "The future prospects of operations research is not solely a question of supply and demand of methods, developed within interdisciplinary environments and involving the aid of specialized scientists who together in a natural way are called upon to solve the problems in question. It is also a question of the attitude of the producers and users of the methods."

and strongly asserted that;

> "Research has always been at the disposal of the rich ones, the strong ones, within society, those who paid."

He concluded his first section by suggesting that;

> "When we are going to look at the future prospects of operations research we have got to start with some of those fundamental working hypotheses on which OR is built and, when opportunity presents itself, ascertain *what* started this very way of thinking within OR and to ascertain the stability of this *what* in a changing world. This is where we shall find the limitations of OR and the weaknesses of our working hypotheses which are of paramount importance to our future assessments."

Professor Jensen then suggested that the central part of 'action-packed operations research' is decision theory. However, he was concerned to emphasise that the reality of decision-making was not always as OR models prescribe. Speaking in particular of politicians he concluded;

> ". . . But these three elementary things, *what do you want, to what extent that is the general attitude of the group*, and the *desire to decide as little as possible*—if it is not going to be too costly not to—we never, or very rarely face within our theory. Our theory is filled with strong men who behind their desks—with the best of intentions but lacking complete knowledge—patch together what the group ought to wish for itself, create the illusion of absolute monarchs or, at the very outside, of Bonus Paters, and who take it for granted that the consequences of the decisions will be carried out—not only within the first phase, but also later on."

His wish to cast mathematical modelling in a functional context led him to turn next to the arguments about optimization:

> ". . . To really use this principle within OR today requires that you stick to rather limited, rather clear tasks not only as far as the operations researcher is concerned, but also as far as his client is concerned. I am afraid that we use operations research far beyond the area for which we can vouch. . . ."

which he followed with a series of related points under the headings of equilibrium, capitalization and communication;

> "As users of mathematics and computation technique we do have a tendency to *put things together in as few units as possible*. . . . We get overwhelmingly happy when we are able to label a single number to a single action. But that is not a good picture of reality, I think, even though it is a convenient part of a rhetorical discussion. . . ."
>
> ". . . It is not only a question of the expert and his immediate customer understanding the situation and its problem. It has also to be easily understood by all the other parties involved in the process if confidence is to be maintained with respect to the application of such methods and with respect to the advisors. But isn't direct assessment dangerous? Yes, it is. But the expert boiling it down into indices is even more dangerous. . . ."

The point was immediately reinforced in a section on "Efficiency, Dehumanisation. Incompleteness."

> ". . . we touch upon one of the perhaps largest deficiencies prevailing at the stage we have reached, and which should have all our attention when we talk about the Future Prospects of OR, namely: the entire OR research and science are influenced by the desire of the Western culture of efficiency. If you at a surprisingly high speed submit the outcome of the measurable results in an easily accessible way without the smoothing, inconsistent, intermediate links which the very complex decision systems previously took advantage of, you have actually essentially contributed to the *dehumanization* of the decision process. The human being plays no part—or only a modest one—in our models and in our herbarium of professional tools.
>
> The role which the human being plays—whether this role is played within a system of *solidarity* or within a system with an *individualism* (not far from selfishness)—is not evident at all, and that role is so important for the final outcome. . . . The *behavioural patterns* do not get a chance. The same goes for the *incompleteness* with respect to the *problem formulation*,

Vol. 5
1977

> *data and knowledge*, both as far as the examining experts and political users are concerned. The fact that the users, when the time comes when they are going to carry out the outcome of the investigation carry it out in an incomplete way, are not sufficiently touched on within our theory. . . ."

We were left with the feeling that much of this presentation was highly unpalatable to many of the audience, and yet Cook [5] has said much the same and received a large response to his ideas from OR workers in the UK. Jensen's general conclusions were harsh;

> "If OR is capable of changing its own research activity, the last part of this century will offer great inspiration. If OR is not capable of this, the next generation will develop without support from us, and OR will, consequently—as so many older disciplines—become a deadweight that has to be overcome."

A theme which Professor Jensen also developed was related to the recognition of conflict and the need to strive towards harmony in society as a major objective relative to that of greater efficiency. This was a topic which was to be returned to later.

Our immediate attention was directed to the contributions of Brehmer, a psychologist, and Holstein, a practising decision analyst. The content of both papers is strongly linked to the importance of decision theory to Operational Research that Jensen had previously discussed. As Brehmer stated early in his paper;

> "Since so many problems can be traced to the covert nature of human judgement, it is not surprising that a large body of research has been directed at finding ways of externalising judgement. This research has taken as its starting point the fact that although a person may not be able to describe how he arrives at his judgements, he is nevertheless able to make them. If we know what judgements a person has made and the information upon which these judgements are based, we are able to build a model which simulates these judgements, and this will give us an insight into the nature of the judgement process. The problem, then, is to find a suitable model."

His attempts to find a suitable model led him to conclude that;

> "Rather than seeking to free (decision-makers) from the biasing influences of emotion and motivation, we should try to find means to augment the cognitive power of decision-makers and to help them overcome the *cognitive* limitations."

It was from this psychological base for Operational Research practice that Holstein took us through the methods he used to help people make real world probability estimates within a decision analysis framework. It is unfortunate that he did not publish a paper for he devoted much of his presentation to an aspect of practice which is rarely discussed—the stages of the interviewing process. He was concerned with aspects such as motivating the client to participate in a decision analysis exercise. His concern for the consultant–client relationship and the techniques he had developed to handle its dynamics

represented an interesting action interpretation of the theoretical framework Brehmer had been anxious to develop. Nevertheless, our own view is that such real world judgements of complex situations are extremely difficult to make, and that it is extremely difficult for us to provide assistance [1]. Perhaps we should pay much more attention to Behavioural Scientists who consider it important to explore the nature of the consultant–client relationships (for example see [7, 13]).

So far in the Congress it was clear that something unusual was going on. As we have indicated, we had heard some of the arguments elsewhere, for example, about the difficulty of providing single 'answers' to complex problems, and about different views of problems from different participants but it seemed that a more sustained review was in progress at the Congress. However, there seemed to be a need for a well-developed and coherent theme. That was to come on the final day, to which we shall now turn. It is convenient to begin with the paper by VS Epstein, Training Director at the General Group of the Swedish Employers' Confederation. After identifying the various forms of what he preferred to call 'worker participation', he stated;

> "Characteristic of the seventies is the tendency of the Unions to use political means to achieve their ends and to prefer laws to agreements. . . .
>
> Behind this legislative activity lies a new union (LO) strategy. The union believes that its objectives cannot be achieved through negotiations. On one issue after another, laws have been demanded to achieve the objectives laid down in 1971. . . ."

Vol. 5
1977

Mr Epstein briefly reviewed the previous situation on the shop floor and the experiments derived from the Norwegian work on autonomous groups. He then turned his main attention to the new legislative changes and their effects. The general content of these laws can be gathered from the Swedish Institute Fact Sheet (FS3h Ohac) issued to us with the Congress documents;

> ". . . The employer is required to consult the employee side prior to taking decisions, nor may he take or implement any decisions before the negotiations are completed. Moreover, the employer is obliged to inform the employee side about all important matters relating to the business—this may also apply to secret undertakings, e.g. as regards trade practices. He will also be obliged to assist the employee side in investigations of the firm which the trade unions think they need in order to take stands on different matters.
>
> Another significant innovation restricts the employer's right to give orders by laying down rules which accord the right of preference in interpreting agreements to the unions. In other words, if the employee side feel that it is not obliged to perform work under the collective agreement, their opinion holds until the dispute is tried and adjudicated."

It is hard to argue that such laws—this being from the 'Act on Employee Participation in Decision-Making' which will come into force in January 1977—do not represent a significant shift in power in Swedish society. Mr Epstein did not avoid the implications for Operational Research;

> "What kind of organisational development is expected? It is obvious that the specialist in operations research must be included in different projects and production groups. However, he may not present his work in the same cryptical and complicated way as heretofore. In the new situation he has to convince not only the management . . . but he has also to convince the representatives of the unions participating in the decision-making group."

We should not assume that these conditions for an Operational Researcher are peculiar to Sweden. The movement towards greater participation in decision-making is widespread [16]. It was interesting to hear the other speaker in this session examine the same issue from a different standpoint. Nygaard explained his own position as a professional worker, a representative of his union on the Research Committee of the Norwegian TUC, and an active participant in helping trade unionists at the shop floor level. His paper was given without notes and no abstract or paper is available. We are therefore unable to quote from it, but only to summarise our impressions. The basic statement was that when he realised that his own work with large-scale computer systems was having major effects on workers without them being taken into account he experienced a significant rethinking of his role. His reaction was to work deliberately with the trade unions to try to help them increase their power to control their own affairs. An example was given of the workings of a Data Agreement in Norway which allows the union side to get in early at the problem formulation stage, and to demand reports in simple language. Thus far the talk was close to that of the employers' representative—in substance if not in tone. Significantly, however, emphasis was also placed on the rights of the unions to control their own educational system for dealing with these new problems, and on the unions' fundamental need to retain working conditions which do not disrupt workers' solidarity—their only source of power. Particularly evident, to us, was the way in which Epstein (although aware of conflict) emphasised harmony between his interests and those of the unions, while Nygaard (although aware of harmony) deliberately emphasised conflict.

Nygaard had a great deal to say directly to OR workers. Perhaps most interesting was his feeling that the freshness of the early days had given way to the staleness of mathematical complexities, yet this new work with the unions offered many new *opportunities* for simple but enormously important (in final effect) uses of our skills. Nygaard finished with two questions;

> "What is the role of OR? What is Ignorance?"

and he made a statement which we thought conveyed very well the cumulative effect of the Invited Speaker Programme;

> "The role of an innocent is no longer possible for an OR worker."

How little room there is left for the innocent OR worker only became completely clear when Sandberg gave the final paper we wish to discuss. So far it would be possible to argue that the reflections on OR had been carried out

essentially from within. However harsh the criticism it was committed. Sandberg's contribution seemed to be a much more dispassionate view of the kind that has been suggested as a scientific orientation [10]. Sandberg opened in the following manner:

> "What is the role of rational analysis in decision-making? What are the functions of operations research, planning methods etc? To answer questions such as these, the rational analysis and the methods must be placed in a social context. It is only against the background of a perspective of society and the organisations that roles and functions can be discussed. But which perspective? Is society a harmonious system, where all members have the same goals and interests? Or is it a framework for conflicts between groups, some of which are rich in resources and dominate the others?"
>
> "Much planning research is concentrated on the question of how planning can contribute to better 'more optimal' or more correct decisions. . . . This can also be expressed such that planning research aims at increasing planning's capability for fulfilling its openly expressed tasks (manifest functions). . . . But . . . planning also has other functions (latent functions) e.g. to legitimate (justify) decisions already taken. . . . The awareness of such latent functions is a prerequisite for a realistic judgement of what can be achieved within an investigatory and planning apparatus as it exists, and without fundamental changes."

The general purpose of the paper was then given;

> "The purpose . . . is to provide a broad, yet concise presentation of the various functions of planning and rational analysis in a societal and organisational context. I shall do so by summarising the functions into two groups and sketching two 'scenarios' for planning's functions in society—colonization and emancipation. . . ."

Vol. 5
1977

There is a danger that the reader may believe that Sandberg's constant references to planning and not OR implies that OR is not involved, it was however clear that Sandberg regarded OR as a significant influence on what he called 'the struggle for the future';

> "The colonization of the future means that today's powerful interests, organised in established institutions, prolong the prevailing situation into the future by anticipating and counteracting crises and threats, accomplishing this inter alia through future studies and planning. Emancipatory future studies and planning, on the other hand, help maintain the possibility of creating alternative futures by means of a broad public participation and change which is in the interests of the majority of the people."

Planning as colonization which he labels as the pessimistic scenario was then characterised in some detail. For example;

> ". . . The functions which prognoses and futures studies have as *self-fulfilling prophecies* is connected to the fact that one of the numerous possible future developments is presented as the probable future behind a wall of calculations and scientificness. . . ."
>
> "Harmony between interests is a presupposition for the application of many planning methods; the planning itself contributes to the definition of situations as harmonious. Planning thought can become the new Latin for elite groups. . . . If social problems are formulated with the aid of the methods of planning and futures studies, interest will be concentrated upon the dimensions which such methods construct. The methods for sharply delimited, quantified concrete problems are more developed than those concerned with

unstructured political problems. . . . Unlike negotiational techniques, analytical techniques aim at investigating the expected consequences of various action programmes for the realisation of common goals, and therefore emphasise a consensus of values and interests. They contribute to the 'construction of consensus', by 'recommending' a frequently implicit mutual perspective."

Planning as emancipation is the total reverse of the above. We do not need to specify all the details given by the speaker, but most importantly we may note;

"The alternatives are presented in political terms and lead to an activation of the democratic process. . . .

Instead of creating harmony and making compromises legitimate, good planning helps to *reveal real conflicts*, both by pointing to alternatives which are based on different fundamental assumptions, and by promoting different interests. . . .

Futures studies and long-range planning become instruments for thorough and *real changes* in society. Instead of passively adapting to developments and solving crises as they appear, people can jointly use futures studies to help anticipate possibilities for change and then use long-range plans to change society and realise desirable futures. . . ."

This notion of constructing models based on different fundamental assumptions and different standpoints is being strongly proposed as the proper method of policy research by at least one major group in the UK [17].

Sandberg then proceeded to his four ideal cases. After noting that the central element in a perspective of reality is its assumptions about harmony and conflict, he pointed out the possibility of a two-level analysis producing four situations;

"These four cases have approximately the following meaning. In manifest conflict the parties are involved in hostilities over incompatible interests. A pseudo-conflict exists when the parties exhibit conflict behaviour in a question where harmony of interests reigns. A latent conflict designates a situation where two parties live in harmony despite the fact that they have irreconcilable interests (this can be due to the fact that the stronger party controls the weaker one). Harmony reigns when neither incompatible interests nor hostile behaviours/attitudes exist."

This led on to a consideration of the functions of any 'rational analysis' in the four cases identified;

". . . Thus, among the people in the empirical reality seen by the observer/researcher is an 'analyst', who is an expert on rational methods of decision-making, and who with respect to harmony and conflict is assumed to exhibit the same behaviour and attitudes as the people in his environment. . . ."

This is clearly an extremely complex issue and it might help the reader to study [21] which fully elaborates his arguments. After dealing in some detail with the role of analysis in these four situations Sandberg identified the implications;

"From a conflict perspective, the analysis cannot contribute to the *elimination* of conflicts (or conflict-solving in the strict sense) until it is an open contribution to a negotiation or a determination of an area of conflict. . . .

> The dominant planning thinking and perspective underlying the utilization of rational analyses in decision-making is a harmony perspective. This provides important insights in many cases, but I believe that it is more fruitful to apply a conflict perspective to planning, analysis and decision-making. . . . A fundamental question is the possibility of a planning which fulfils either the one function or the other. What are the conditions for an emancipatory planning?"

This question was the topic of the final section of the paper;

> ". . . In a society where there exist fundamental contradictions, the most important condition for emancipatory planning is that the plans explicitly consider such contradictions. It is not sufficient, however, merely to write about them and present alternatives. It is also necessary that the dominated party have sufficient resources to examine critically the commission reports presented and to conduct futures studies and planning from its own perspective and interest, as well as to carry out these plans—and eliminate the dominance.
>
> Established planning thought and its reflection in handbooks and textbooks about planning often state as its purposes much of that which we here have called the emancipatory functions, which are presented as unproblematic. This is due to the fact that they have a harmony perspective, and therefore do not discuss the organisational, and even less so the societal conditions for a planning with such functions. The analysis of methodology, objectivity and scientificness in futures studies and planning is quite important, but it is primarily the practical context in which they occur which determines their function."

The extent to which many writers on Operational Research have established a need to question who OR works for and the extent to which the social context of analysis influences the ways in which reality is officially defined needs to be acknowledged. Notably Rosenhead [20] and Hales [14] have conducted their own research and writing from a political science or power analysis; in addition Bevan and Bryer [3] recently specifically took up the relevance of power in decision-making.

ANALYSIS

The Invited Speakers' Programme apparently was intended to be a major element of EURO II as the papers were given in plenary sessions or the main conference chambers during parallel sessions. However, many papers were published separately from the Congress Proceedings, which explains in part our desire to present the above review in an easily accessible place. Nevertheless, the main reason for this paper is to reveal the contrast between the content of the Invited Papers and that of the rest of the ('Official') Congress. We shall argue later that this must have major implications for the practice of OR.

The decision-makers invited to speak at EURO II emphasised the complexity of the problems that we are increasingly having to face. They also stressed the very self-aware OR which they considered useful in such difficult circumstances. Technical expertise was required, but also an understanding of the problematic bases and social context of the technical work.

The presentations of the OR workers invited to speak were remarkable for their interdisciplinary approaches, different speakers drawing upon various combinations of Political Science, Sociology, Psychology and Philosophy. Despite this variety the papers all shared one characteristic—*introspection*. All of the speakers were willing to think hard about their own studies in order to discover good paradigms for future OR practice. In addition these speakers (and, of course, Brehmer) acknowledged the uncertainties of human judgement in complex situations. Finally, all seemed to recognise clearly the issues of power relationships and conflicts in organisations and society. Related to this, we suspect, was their reluctance to use Systems Frameworks. They seemed to prefer an Action Framework in which a specific interest group was unambiguously identified as the client for the analytical work.[3] In short, they adopted a framework which asserts that reality cannot be objectively defined in complex social situations, or as Kelly [15] has put it;

> ". . . the events we face . . . are subject to as great a variety of constructions as our wits will enable us to construe."

Naturally enough in situations of power differences and conflict, the interests of the parties have a great effect on what each makes of the circumstances [11]. In our opinion the features noted above—an interdisciplinary approach, acknowledgement of the difficulties of making judgements, recognition of conflicts and power differences, and wide-ranging introspection—are precisely what is required to generate the level of self-awareness considered necessary by the experienced decision-makers at EURO II.

Unfortunately, the feeling we had at the Congress—that there was a major difference between the invited papers and the rest of the 'Official' papers—was confirmed by a detailed analysis. Of the approximately 150 'Official' contributions, only a minute proportion could be said to significantly share the characteristics of the Invited Papers. The overwhelming majority of these papers followed a Systems or Mathematics approach, solely in relation to technical problems. This, clearly, was not necessarily due to the nature of the programme. Production and Inventory and Transportation, for example, are indeed subjects where a purely technical paper is quite legitimate, but they are also subjects in which technical studies can produce major impacts in the real world.

[3] These two frameworks are discussed by David Silverman [22] from a sociological standpoint. The Systems perspective is well known amongst OR workers. However, its implications for the way in which it implicitly influences the form of analysis that is conducted is rarely discussed in an OR context. Recently Bryer and Kistruck [5] have produced a working paper which attempts to do this. Generally speaking we would expect to see a correlation between the Systems perspective and a dominance of Economics and Mathematics in OR.

IMPLICATIONS

We do not claim that this is an unbiased account of EURO II. It is clearly prejudiced by our own previously stated views on the worrying aspects of much OR [10, 12]. Our view is perhaps best summarised by Collins' [8] account of the effects of the development of a proper social science;

> "The applied fields, finally, should also be changed. For the most part, they will have to become less pretentious. Certainly, they will have to fight themselves free of the quackery that prevails now, as value choices are hidden under displays of technical jargon or collections of facts—which never in fact speak for themselves. A truly serious applied social science cannot attempt to hide value choices, but must make them explicit. . . ."

Nevertheless in spite of our own bias it would seem hard to argue that a gap between the Invited Papers and 'Official' Proceedings does not, in fact, exist.

Should the above have any implications for the OR practitioner? We cannot help concluding that EURO II was yet another public demonstration of the 'technically bounded perspective' which pervades so many OR conferences, journals and education courses. Can we avoid the consequence that such public demonstrations colour the views that potential clients, students and the general public have of OR? Furthermore, is it possible to avoid concluding from EURO II that there is a massive gap between our aspirations to be interdisciplinary practitioners (see the Invited Speakers) and the reality (see the Proceedings)? Finally, could not our restricted perspective be at the root of our many so-called 'Implementation Problems'? The reader may disagree, but we heard people (other than the Invited Speakers) at EURO II saying;

> "Of course OR does not work in the way published papers suggest"

or;

> "The politics of organisational life are crucial to achieving effective implementation."

As always, however, this was said backstage. Even if such views are widely held and applied in OR we might take note of what Churchman [6] said in 1967 in relation to "wicked problems";

> "To date, operations research and management science have been largely indifferent to the morality of the profession, perhaps because the profession has not yet taken itself seriously. That the profession has a moral problem, nonetheless, there can be no doubt. It might make us look more mature if we began to discuss it."

The problems our society faces will, we believe, require as much relevant technical expertise as science-based analysts can provide. Yet it is beginning to be widely appreciated that technical expertise alone is insufficient—or even

Vol. 5
1977

counter-productive. We have noted previously the interest that this matter has aroused in the UK. This concern is not, of course, restricted to the UK, in North America, for example, Churchman [6] and the Wharton School Group [19] are well-known proponents of similar views.

It would seem to us that we all owe a debt to the Scandinavian speakers for the cogent argument that they severally and jointly presented. Surely they demonstrated convincingly that if we wish to assist effectively with complex social problems (and we have expressed that desire often enough) then we must develop a much greater ability to reflect on our own studies. If EURO II is any guide this means a fundamental reorientation of much OR practice. If this reorientation comes about, EURO III will probably demonstrate a far greater impact of our work on the real world.

REFERENCES

1. BERRESFORD A & DANDO MR (in press) Operational research for strategic decision-making; the role of world-view. *Opl. Res. Q.*
2. BISHOP BC (1972) A contribution to a discussion on the methodology of operational research. *Opl. Res. Q.* **23**, 251–61.
3. BEVAN RG & BRYER RA (1976) The Aspirations and the Practice of OR. Presented to the OR Society Conference at the University of Swansea.
4. BOWEN KC (1974) Helping with Decisions on Complicated Systems. Paper presented to the Second International Conference on Systems and Informatics, Mexico.
5. BRYER RA & KISTRUCK R (1976) Systems Theory and Social Science: the Problems of Applying Systems Concepts to Social Organizations. School of Industrial and Business Studies, University of Warwick Working Paper.
6. CHURCHMAN CW (1967) Guest editorial on wicked problems. *Mgmt Sci.* (B) **14**(4), 141–42.
7. CLARK PA (1972) *Action Research and Organizational Change*. Harper Row, London.
8. COLLINS R (1975) *Conflict Sociology: Toward an Explanatory Science*. Academic Press, New York.
9. COOK SL (1973) Operational research, social well-being and the zero growth concept. *Omega* **1**, 647–667.
10. DANDO MR, DEFRENNE A & SHARP RG (1977) Could OR be a Science? *Omega* **5**(1), 89–92.
11. EDEN C (1975) An Exploration of the Role of a Decision Analyst, School of Management, University of Bath; and also Operational Research and Organization Development, presented at EURO II.
12. EDEN C and HARRIS J (1975) *Management Decision and Decision Analysis*. Macmillan, London.
13. EDEN C & SIMS D (1976) Problem definition between consultant and client. Centre for the Study of Organisational Change and Development, University of Bath.
14. HALES M (1974) Management science and the 'second industrial revolution'. *Radical Sci. Jl* **1**, 5–28.
15. KELLY GA (1955) *The Psychology of Personal Constructs*. Norton.
16. LUCAS JR (1976) *Democracy and Participation*. Penguin, London.
17. MCLEAN JM (1976) The Simplicity and Adequacy of Models—a Case for Structural Pluralism. Science Policy Research Unit, University of Sussex.
18. MITCHELL GH (1973) The State of Research in OR. *Opl. Res. Q.* **24**(1), 3–9.
19. OZBEKHAN H (1974) Thoughts in the emerging methodology of planning, in *Systems and Management Annual* (Ed. ACKOFF RL). Petrocelli Books, New York.

20. ROSENHEAD J (1977) *The Technology of Political Control*. Penguin, London.
21. SANDBERG A (1976) *The Limits to Democratic Planning; Knowledge, Power and Methods in the Struggle for the Future*. Liber Forlag, Sweden.
22. SILVERMAN D (1970) *The Theory of Organizations*. Heinemann.

ADDRESS FOR CORRESPONDENCE: *Dr Malcolm Dando, The University of Sussex, Operational Research, Mantell Building, Falmer, Brighton BN1 9RF, UK.*
Dr Colin Eden, The University of Bath, BA1 7AY, UK.

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

Maintaining the Calibration of an Information System

PHILLIP G CARLSON

Emory University, Atlanta, USA

(*Received March* 1976; *in revised form November* 1976)

The design and operation of an information system requires assumptions on the values of parameters which characterize the behavior of pertinent random processes. If these parameters remain constant over time at design values, system performance will be as specified. Otherwise, performance will vary from design and cost and service penalties will arise. A control chart procedure is introduced for monitoring system parameters and introducing appropriate action to reset them to design values as indicated. The control system is set so that the expected cost of monitoring is a minimum. A table of control system settings over a broad set of *a priori* assumptions is provided. Examples describing the procedure and use of the table are given.

INTRODUCTION

IN THE design of a management information system assumptions are made on the behavior of pertinent random processes. The values of their parameters provide, by design, a basis for predicting the performance of the information system. Should the parameters vary from the preset values, performance can vary from design and a penalty may be incurred.

Assuming that system calibration is to be maintained, the procedures described here provide a basis for deciding when to take how much action to maintain that calibration.

It is the objective to develop criteria for detecting when a system is off calibration and for taking proper action to reset the affected parameters. An information system does not perform only because a given situation has been described based upon a set of plausible assumptions. It must be made to perform and in the design of such systems means must be provided for forcing certain relationships to hold. One must build in independent variables which can be manipulated, as well as rules for their manipulation, to ensure the preservation of preset parameter values and to maintain proper system performance. It is generally known that n manipulable input variables are necessary and sufficient for the control of n output variables. The case where $n = 1$ will be considered here.

When the system is on calibration, specified performance can be expected; if one or more parameters have shifted from their prescribed values, the system will likely not deliver specified performance. Output can vary from design in two ways: first, the system can provide inadequate performance with a resultant service penalty to the user; second, the system can provide too good performance, with an unjustified cost penalty to the user.

To maintain the desired performance, and take restorative action as indicated, we hypothecate the availability of manipulable variables which can be used to effect reset action without otherwise affecting the system. We would expect to anticipate performance breakdowns in time to avoid most actual failures, providing a path between the extremes and encountering one of them only occasionally.

A PROPOSED APPROACH

Consider an information system. Let B be a parameter whose value affects system performance. Usually B will depend on random processes which are part of the system. Let the set value of B be B_0. When $B = B_0$, the system will perform as specified. If B is much greater than B_0, performance will deteriorate and penalties will be incurred; if B is much less than B_0, performance will improve, but also at a penalty.

Vol. 5
1977

We shall be concerned with a control system for maintaining the value of B at or near B_0. If the value of B changes we shall want to discover it promptly, and, using a suitable manipulable variable, effect a timely restoration of B to the value B_0. Further we shall be concerned with the determination of an optimal value of B, namely the value B_0, both from the viewpoint of the operation of the information system and the control system. We shall use an example to develop the procedure.

Example: an information system

To demonstrate the development of the control system we shall look at a typical reorder point inventory management system which triggers a sequence of orders which themselves become demands on a production facility.

Consider a non-seasonal product line A comprising n items $A_1, A_2, \ldots, A_n$ which is sold to customers on demand from a warehouse stock. Let the warehouse be supplied from a single source factory which makes only product line A and ships its entire output to the warehouse. Let D_i be the daily demand for item A_i in standard production days. Let D_i be approximately normally distributed with mean Δ_i and standard deviation σ_i. Then if the replenishment or lead time λ is a constant λ_0 days the reorder point for item A_i is the familiar

$$\mathrm{ROP}_i = \lambda_0 \Delta_i + k_i \sqrt{\lambda_0}\, \sigma_i \qquad \text{(I)}$$

where k_i is the number of units of service stock, that is, the number of standard

deviations of demand during a lead time, associated with the i^{th} item. When the stock on hand or on order for item A_i drops below ROP_i, a lot of size Q_i is ordered from the factory.

Equation (1) provides a replenishment level for a product A_i as a function of the lead time demand, where k_i is a measure of customer service. If the demand distribution and the lead time are as specified, with production geared to meet average demand, then customer service for the products A_i will be as specified by the appropriate k_i. In practice Δ_i is replaced with $\bar{D}_i$ in equation (1) where $\bar{D}_i$ is a frequently updated sample estimator of Δ_i and where $E\bar{D}_i = \Delta_i$.

The aggregate daily demand for all items A_i is D, with mean Δ and variance σ^2. The aggregate service stock is the sum of all the service stocks. Then if k is chosen such that

$$k\sqrt{\lambda_0}\sigma = \sqrt{\lambda_0} \sum_{i=1}^{n} k_i\sigma_i$$

we can think of an equivalent aggregate reorder point

$$\text{ROP} = \lambda_0\Delta + k\sqrt{\lambda_0}\sigma \qquad (1')$$

where $k\sqrt{\lambda_0}\sigma$ is the entire service stock.

The factory production rate ρ is assumed constant, or, more generally, an independent manipulable variable. The variable ρ can be manipulated in two distinct ways: one is by varying the size of the production force; another is by varying the length of the work week.[1]

If the overall average daily demand, Δ, shifts upward while the factory production rate, ρ, remains constant, the true lead time, λ, must increase and the value of lead time, λ_0, in equations (1, 1′) will not be correct in that it will be too small. This will be realized in a buildup of production backlog, and stock will not be available to meet customer demands at the objective service level. To restore the lead time value, λ_0, and objective customer service, action is required to reduce the backlog by an increase in the production rate, ρ.

Similarly, if Δ shifts downward while ρ remains constant, the lead time will decrease and production backlog will be reduced. If action is not taken to restore the lead time to the value λ_0, inventories will increase inordinately and customer service will improve to an unjustifiedly high and costly level.

Neither situation is desirable.

A CONTROL SYSTEM

One way to maintain the lead time parameter value near a prescribed value, λ_0, is to maintain production backlog B, in standard production days, near a

[1] Some may argue that shortening the work week is often not an option open to management. In cases where this is a significant consideration, biasing the result given below, at an additional cost, can all but preclude the necessity of that action.

certain value, say B_0, on the basis that if backlog is essentially constant, then the same can be said for order lead time.

Backlog is considered to be the cumulative difference between daily production orders (essentially demand) and orders put into production (essentially amount produced). Assume that

$$B_t = B_0 + \sum_{j=0}^{t} (D_j - \rho_j)$$

is the backlog at time t, that is, t days after the most recent reset to B_0. The time difference of B_t is $\dot{B} = D - \rho$. Hence, under design conditions the expected rate of change of backlog, $E\dot{B}$ is zero. Also, Var $\dot{B}$ = Var $D = \sigma^2$ since the production rate ρ is assumed constant.

The procedure is to observe production backlog daily and use the sequence of observations to test hypotheses concerning the average time rate of change of the backlog, that is concerning $E\dot{B}$. To ease notation say $E\dot{B} = \theta$. We shall start by testing:

$$H_0\colon \theta = -\epsilon < 0 \text{ vs. } H_1\colon \theta = \epsilon$$

where $P(H_0|H_1) = \beta$ and $P(H_1|H_0) = \alpha$. That is, we use the sequential observations on B to test the hypotheses that production backlog is decreasing at rate ϵ as opposed to the alternative that backlog is increasing at rate ϵ. If the test terminates with the acceptance of H_0 we say that λ is not constant at the design value but is decreasing. We restore the backlog to its design level, B_0, and thus the lead time parameter to its design value, by increasing the value of B by an amount ΔB where ΔB is to be determined. Alternatively, if the test terminates with the acceptance of H_1 we say that λ is not constant at the design value but is increasing. The proposed action is to decrease B by an amount ΔB. (Implied is $\beta = \alpha$) B_0 is related to λ_0 by $B_0 = \lambda_0 + \Delta B$. Further, once reset has been indicated a new test is started.

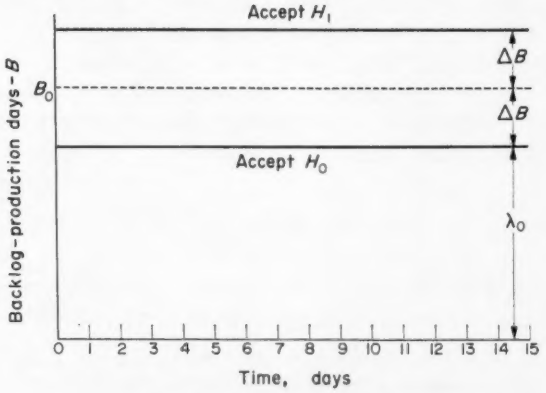


FIG. 1. *Control action chart.*

Vol. 5
1977

Again, decreasing B by an amount ΔB requires the adding of ΔB production days (overtime!) and increasing B by a similar amount requires the omitting of ΔB production days (undertime!).

Performing the test is highly amenable to the use of control chart techniques and a convenient way of doing it is, starting from $B = B_0$, to plot B daily on a chart such as that given in Fig. 1. The nominal line on the chart is at $B = B_0$, the lower limit is at $B = B_0 - \Delta B$, and the upper limit is at $B = B_0 + \Delta B$. The chart starts at $t = 0$, immediately after the latest reset decision. The first time that B meets or goes above (below) the upper (lower) limit, the hypothesis H_1 (H_0) is accepted, and action is taken to decrease (increase) B by an amount ΔB, thus restoring B to the value B_0. With this decision the test is terminated and $B = B_0$ provides the startup observation for the next test.[2]

The main task in setting up the control chart is determining the economic value of ΔB.

THE ANALYSIS OF COSTS

In an operating situation action will be taken occasionally over a planning period, say T days, to maintain an information system on calibration. Such actions are considered independent. It is assumed that any cost of action is incurred during the period in which the decision is made, and that this cost is part of the overall cost of making decisions.

There are two costs which are involved in operating the control system, and their sum over the period T is an appropriate performance measure. One of these is the cost of making a decision, that is, the cost of a reset action, of changing backlog by an amount ΔB. The other is the cost of additional inventory required in the information system in support of the control system, by virtue of increasing the order lead time from λ_0 to $\lambda_0 + \Delta B$ days.

First consider $C_i(\theta,A)$, $i = 0,1$ where $A = (1-\alpha)/\alpha$. $C_i(\theta,A)$ is the cost of accepting H_i when the true rate of change of backlog is θ. $C_0(\theta,A)$ is the cost of taking action to reset, that is, increase, the backlog an amount ΔB when H_0 has been accepted and θ is the true rate of change of B. $C_1(\theta,A)$ is the corresponding cost when H_1 has been accepted.

The cost functions imputed are shown in Fig. 2 along with their equations. In particular the cost of a decision is proportional to the amount, ΔB, of reset action and a scale factor M which would be determined from local conditions. The costs of incorrect action are seen to be quite high, while a correct action incurs zero cost.

The second cost, $C_1(\epsilon,A)$ is the cost of the additional inventory necessary to operate the control system, since B_0 is set at ΔB production days more than

[2] If reset actions are frequent and primarily in the same direction then a permanent shift in the production rate may be indicated.

the nominal backlog, or in terms of lead time days, ΔB days greater than λ_0.

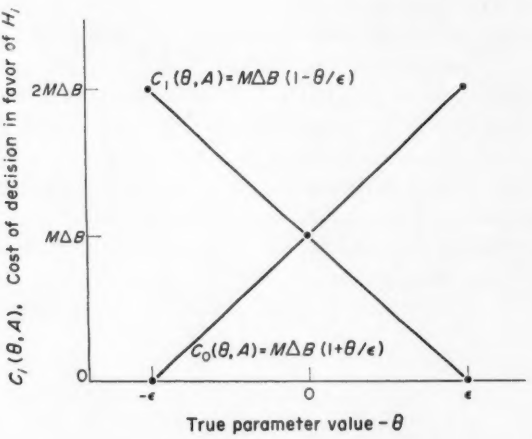


FIG. 2. *Imputed cost of a decision.*

Vol. 5
1977

The development of an expression for the increase in inventory follows.

From equation (1′) it is seen that the change in service stock, ΔI, in the warehouse resulting from increasing the order lead time from λ_0 to $\lambda_0 + \Delta B$ days is

$$\Delta I = k\sigma\,[\sqrt{\lambda_0 + \Delta B} - \sqrt{\lambda_0}] \sim k\sigma\Delta B/2\sqrt{\lambda_0} = k\sigma^3 \log A/4\epsilon\sqrt{\lambda_0}, \text{ where } \Delta B \ll \lambda_0, \quad (3)$$

from Wald's result[3] [4] that $\Delta B = \sigma^2 \log A/2\epsilon$.

If the cost of carrying a day of production in inventory for one day is C_p, then the cost of carrying ΔI for the T day planning period is

$$C_I(\epsilon, A) = C_p T k\sigma^3 \log A/4\epsilon\sqrt{\lambda_0}. \quad (4)$$

To obtain the cost of making decisions over the period T, the number expected to be made during T days must be obtained. Toward this, $En(\theta)$ is the average sample number, that is, number of days to a decision, given that θ is the true parameter value.

The ratio $T/En(\theta)$ will be used as an approximation to $E[T/n(\theta)]$ which is the expected number of decisions made during a period of length T. Also, $[T/En(\theta)]\cdot L(\theta,A)$ and $[T/En(\theta)][1-L(\theta,A)]$ are estimates of the average numbers of decisions made in favor of H_0 and H_1 respectively during a period of length T, where $L(\theta,A)$ is the probability of accepting H_0 when θ is the true parameter value given A. Then, the expected cost of making decisions over the period T, given θ, is

[3] The appendix outlines the essentials of sequential analysis necessary for tracking the development.

$$C_D(\theta,A) = [T/En(\theta)]\,\{C_0(\theta,A)\,L(\theta,A) + C_1(\theta,A)\,[1 - L(\theta,A)]\}$$

and, if the inventory carrying cost is added, the total period cost becomes

$$C_T(\theta,\epsilon,A) = C_D(\theta,A) + C_I(\epsilon,A). \tag{5}$$

We are interested in the value of A, and hence of α, that will lead to a minimum total period cost, subject to an assigned value of ϵ and an *a priori* distribution of θ.

A CLASS OF PROBLEMS

Here we outline a class of representative situations, develop and tabulate minimum cost solutions, and show how they can be used to set up and calculate the properties of a control system.

Assume here that the parameter θ can take on only three values, namely, $\theta = -\epsilon$, $\theta = 0$, and $\theta = \epsilon$, that backlog is decreasing at rate ϵ, remaining constant, or increasing at rate ϵ. Hence we have

$$\begin{aligned} C_D(\theta,A) &= 2M\epsilon T/(A-1) && \text{for } \theta = -\epsilon; \\ &= 2M\epsilon T/\log A && \theta = 0; \\ &= 2M\epsilon T/(A-1) && \theta = \epsilon. \end{aligned}$$

Consider a corresponding family of *a priori* distributions of θ of the form

$$P(\theta = -\epsilon) = P(\theta = \epsilon) = \gamma/2, \text{ and } P(\theta = 0) = 1 - \gamma,\ 0 \leq \gamma \leq 1,$$

where γ is the *a priori* factor.

Further let $R = 2MT$ and $S = kC_pT\sigma^3/(4\sqrt{\lambda_0})$, and equation (5) becomes

$$C_T(A) = (1 - \gamma)\,R\epsilon/\log A + \gamma R\epsilon/(A - 1) + S \log A/\epsilon. \tag{6}$$

This is a typical cost balance situation, given ϵ and γ, as a function of A: the first two terms decrease with A while the third term increases with A. For assigned values of ϵ and γ, and values of R and S obtained from the particular process, the optimal solutions of (6) can be obtained from Table 1.

Table 1 is set up as follows: The columns are for values of the *a priori* factor γ from 0·00 to 1·00 in increments of 0·10 and the rows are for values of α from 0·010 to 0·100 in increments of 0·01. In the body of the table corresponding to each pair (γ,α) are three entries: the top one is the action limit parameter $\sqrt{R/S\epsilon}$, the next is the total cost parameter, $C_T/R\epsilon$, and the last is the average sample number parameter $\epsilon^2 En/\sigma^2$.

TABLE 1. LIMIT PARAMETER: COST PARAMETER: SAMPLE NUMBER *a priori* FACTOR

RISK	0·00	0·10	0·20	0·30	0·40	0·50	0·60	0·70	0·80	0·90	1·00
0·010	4·595	4·786	5·003	5·253	5·544	5·889	6·308	6·832	7·513	8·448	9·849
	0·435	0·397	0·360	0·322	0·284	0·246	0·209	0·171	0·133	0·095	0·058
	5·279	4·976	4·673	4·371	4·068	3·765	3·462	3·160	2·857	2·554	2·252
0·020	3·892	4·031	4·186	4·360	4·559	4·787	5·053	5·369	5·753	6·233	6·857
	0·514	0·473	0·432	0·391	0·350	0·309	0·268	0·227	0·186	0·145	0·104
	3·787	3·595	3·403	3·211	3·019	2·827	2·635	2·444	2·252	2·060	1·868
0·030	3·476	3·586	3·706	3·840	3·990	4·158	4·349	4·570	4·828	5·136	5·510
	0·575	0·532	0·490	0·447	0·404	0·361	0·318	0·275	0·232	0·189	0·146
	3·021	2·882	2·743	2·605	2·466	2·327	2·189	2·050	1·911	1·772	1·634
0·040	3·178	3·268	3·366	3·473	3·591	3·722	3·868	4·034	4·222	4·440	4·695
	0·629	0·585	0·541	0·497	0·453	0·408	0·364	0·320	0·276	0·232	0·188
	2·525	2·419	2·312	2·206	2·100	1·993	1·887	1·781	1·675	1·568	1·462
0·050	2·944	3·020	3·101	3·189	3·285	3·390	3·507	3·636	3·780	3·943	4·129
	0·679	0·634	0·589	0·544	0·499	0·454	0·409	0·364	0·318	0·273	0·228
	2·167	2·083	1·999	1·915	1·830	1·746	1·662	1·578	1·493	1·409	1·325
0·060	2·752	2·815	2·884	2·958	3·038	3·124	3·219	3·322	3·436	3·563	3·705
	0·727	0·681	0·635	0·589	0·544	0·498	0·452	0·406	0·360	0·314	0·269
	1·893	1·825	1·756	1·688	1·620	1·552	1·483	1·415	1·347	1·279	1·211
0·070	2·587	2·642	2·700	2·763	2·830	2·902	2·980	3·065	3·157	3·259	3·371
	0·773	0·727	0·680	0·634	0·588	0·541	0·495	0·448	0·402	0·355	0·309
	1·673	1·617	1·561	1·505	1·449	1·393	1·336	1·280	1·224	1·168	1·112
0·080	2·442	2·490	2·540	2·594	2·651	2·712	2·777	2·848	2·924	3·006	3·096
	0·819	0·772	0·725	0·678	0·631	0·584	0·538	0·491	0·444	0·397	0·350
	1·491	1·445	1·398	1·352	1·305	1·259	1·212	1·165	1·119	1·072	1·026
0·090	2·314	2·355	2·399	2·445	2·494	2·546	2·601	2·660	2·724	2·792	2·865
	0·864	0·817	0·770	0·723	0·675	0·628	0·581	0·533	0·486	0·439	0·392
	1·338	1·299	1·260	1·221	1·182	1·143	1·104	1·065	1·027	0·988	0·949
0·100	2·197	2·233	2·271	2·311	2·354	2·398	2·445	2·495	2·549	2·606	2·667
	0·910	0·863	0·815	0·767	0·720	0·672	0·624	0·577	0·529	0·482	0·434
	1·207	1·174	1·141	1·109	1·076	1·043	1·010	0·977	0·945	0·912	0·879

To use the table first determine R and S from local conditions and impute ϵ and γ. Then calculate $\sqrt{R/S\epsilon}$ and find the nearest value in the column of the imputed value of γ. This determines the value of α (in the left column) to the nearest 0·010. Then all desired parameter values can be obtained.

Assume the following set of parameter values to obtain when a company decides to install a control system to maintain the calibration of its inventory management system: a planning period of $T = 250$ operating days, an objective order lead time of $\lambda_0 = 25$ days, a carrying rate of $C_P =$ \$100 per day per production day in inventory. The standard deviation of total daily demand is $\sigma = 0{\cdot}1$ production day, the overall service factor is $k = 2$, and the scale factor M is \$1000. Using these values in the expressions for R and S we have $R =$ 500,000 and $S = 2{\cdot}5$.

Case I. Management feels that the chance that backlog will remain constant during the ensuing period is about 0·3 and the chances are about even that it

Vol. 5
1977

will increase or decrease at a rate of 0·01 production day per day. Determine a minimum cost backlog control system to maintain the inventory system on calibration in the light of this conjecture.

First, this is interpreted as $\epsilon = 0{\cdot}01$, $P(\theta = 0) = 0{\cdot}3$ and $P(\theta = -0{\cdot}01) = P(\theta = 0{\cdot}01) = 0{\cdot}35$ or $\gamma = 0{\cdot}7$. Now $R/S = 200{,}000$ and $\sqrt{R/S} = 447{\cdot}2$ which makes 4·472 the value of $\sqrt{R/S\epsilon}$. The nearest value in the column headed $\gamma = 0{\cdot}7$ is 4·570 which corresponds to an α of 0·03 (we shall not interpolate here). Then the value of $C_T/R\epsilon$ is 0·275 indicating a total period cost of \$5000 (0·275) or \$1375. The value of $\epsilon^2 En/\sigma^2$ is 2·050 which gives an En of 205 days to a decision (For $\alpha = 0{\cdot}03$, the value of A is 32·333).

The distance from the B_0 value to the upper (or lower) limit line is $\Delta B = \sigma^2 \log A/(2\epsilon) = 1{\cdot}738$ days and the value of B_0 is $\lambda_0 + \Delta B = 26{\cdot}738$ days. The amount of additional inventory required by the inventory management system is $\Delta I = k\sigma(\Delta B)/(2\sqrt{\lambda_0}) = 0{\cdot}035$ production days. The decision cost functions are, respectively, $C_0(\theta,32{\cdot}33) = 173{,}800\theta + 1738$ and $C_1(\theta,32{\cdot}33) = -173{,}800\theta + 1738$ as calculated from the expressions in Fig. 2.

Summarizing, if the assumptions made are truly representative of the operating situation one can expect a decision and a reset action about once every 205 operating days or about 4 actions in 5 years. The total period cost of making decisions and carrying the additional inventory increment is \$1375 per 250 operating days.

The set point for the control chart is at $B_0 = \lambda_0 + \Delta B$ or 26·738 days which is also the new lead time value to be used in the inventory management system. The additional inventory required to accommodate the increased lead time is 0·035 production days.

Case II. Management feels that in the ensuing period demand will remain essentially constant and there is little chance it could change by 0·005 production day per day. In the light of this, design a minimum cost control system to keep the inventory management system on calibration.

This is interpreted to mean that $\epsilon = 0{\cdot}005$, that $P(\theta = -0{\cdot}005) = P(\theta = 0{\cdot}005) = 0{\cdot}05$, and that $P(\theta = 0) = 0{\cdot}9$ or that the *a priori* parameter value is 0·1 Then $\sqrt{R/S\epsilon} = 2{\cdot}236$. Entering the column headed $\gamma = 0{\cdot}10$ the nearest value is 2·233 which corresponds to an α of 0·10 and an A of 9. Also (not interpolating) $C_T/R\epsilon = 0{\cdot}863$, and $C_T = \$2158$ per period. Further $\epsilon^2 En/\sigma^2 = 1{\cdot}174$ and $En = 470$ days. The parameter ΔB has the value 2·197 production days, ΔI is 0·044 production days, and B_0 is 27·197 days. The decision cost functions are, respectively, $C_0(\theta,9) = 219{,}700\theta + 2197$ and $C_1(\theta,9) = -219{,}700\theta + 2197$.

Summarizing, if the assumptions are truly representative of the situation, a decision and the accompanying reset action can be expected about once every 470 operating days or about once in two years. The probability of accepting H_0 when H_1 is true, or vice versa, is 0·100 or about one chance in 10. The chances of accepting either H_0 or H_1 when the mean backlog is truly constant ($\theta = 0$)

are even. The total period cost of making decisions and carrying the additional inventory increment is $2158.

The set point for the control chart is 27·197 days which is also the new lead time for the inventory system. The additional inventory required to accommodate the increased lead time is 0·044 production days.

ACKNOWLEDGEMENT

The author wishes to acknowledge the generous counsel of Sebastian B Littauer.

REFERENCES

1. EHRENFELD S & LITTAUER SB (1964) *Introduction to Statistical Method.* McGraw-Hill, New York.
2. HOPEMAN RJ (1976) *Production: Concepts, Analysis, and Control.* Merrill, Columbus.
3. WAGNER HM (1970) *Principles of Management Science.* Prentice-Hall, Englewood Cliffs, NJ.
4. WALD A (1947) *Sequential Analysis.* Wiley, New York.

APPENDIX—SEQUENTIAL ANALYSIS

This summary has been adapted from Wald [4].

Let X be a normal random variable with unknown mean θ and known variance σ^2. Consider a test of hypothesis H_0: $\theta=\theta_0$ vs. H_1: $\theta=\theta_1$ where $P(H_1|H_0) = \alpha$ and $P(H_0|H_1) = \beta$ are the risks of Type I and Type II error. Under H_i, $i=0, 1$ we have

$$f(x; \theta_i) = (1/\sqrt{2\pi}\sigma) \exp -\tfrac{1}{2} [(x - \theta_i)/\sigma]^2, \quad -\infty < x < \infty.$$

Consider a sequence of observations $x_1, x_2, \ldots, x_i, \ldots$ and the associated $z_1, z_2, \ldots, z_i, \ldots$ where $z_i = \log[f(x_i; \theta_1)/f(x_i; \theta_0)]$. At each stage i observe the cumulative $z_1 + z_2 + \ldots + z_i$, then where $A=(1 - \beta)/\alpha$ and $B=\beta/(1-\alpha)$, if

(1) $\log B < z_1 + z_2 + \ldots + z_i < \log A$, take another observation;

(2) $\log B > z_1 + z_2 + \ldots + z_i$, accept H_0, terminating the test;

(3) $\log A < z_1 + z_2 + \ldots + z_i$, accept H_1, terminating the test.

The test continues until condition (2) or (3) appears for the first time.

The operating characteristic function of the test is given by

$$L(\theta) = (A^{h(\theta)} - 1)/(A^{h(\theta)} - B^{h(\theta)}) \text{ where } h(\theta) = (\theta_1 + \theta_0 - 2\theta)/(\theta_1 - \theta_0).$$

The average number of observations to termination of the test is given by

$E_n(\theta)=[L(\theta)\log B + (1 - L(\theta))\log A]/E(\theta,z)$ where $E(\theta,z) = [2(\theta_1 - \theta_0)\theta + \theta_0^2 - \theta_1^2]/2\sigma^2$.

The slope of the reference line for the graphical procedure is $(\theta_0 + \theta_1)/2$ and the distances from the reference line to the action lines are $\sigma^2\log B/(\theta_1 - \theta_0)$ and $\sigma^2\log A/(\theta_1 - \theta_0)$ for accepting H_0 and H_1, respectively.

ADDRESS FOR CORRESPONDENCE: *Professor Phillip G Carlson, School of Business Administration, Emory University, Atlanta, GA* 30322, *USA.*

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

Human Dynamics: A New Approach to Human Behavior

RA THIÉTART

Ecole Supérieure des Sciences Economiques et Commerciales, Cergy, France

(*Received August 1976; in revised form January 1977*)

This paper presents a new approach to analyze human behavior and more precisely human performance. It calls upon a system dynamics methodology to develop a general framework for looking at human beings. As an illustration of the approach, one subsystem is described. The global human system is also briefly presented. The contribution of such an approach is illustrated with an example dealing with the study of a worker performance subject to different types of control.

1.5
77

INTRODUCTION

BECAUSE of the high complexity of models of human behavior dealing with many variables and time-varying components, it is difficult to derive any significant inferences from their analysis. The difficulty in grasping, at the same time, many different elements of the model, which are changing in a dynamic fashion, obliges one to find new ways for looking at and analyzing human behavior.

An approach emphasizing the existence of a human behavior system, stressing the mutual influences of variables and their time-varying characteristics, seems worth adopting. However, to be able to take full advantage of this broader approach, one must place more reliance upon mathematical modelling and simulation. System simulation, which embodies the methodology of the present approach, has already been extensively used in engineering and industrial fields. However, because it is difficult "to crack the system oyster, to locate the psychological pearls", as Abelson [1, p. 279] points out, no comprehensive research has yet been undertaken to build a human system simulation model. Human Dynamics is the first comprehensive attempt.

Human Dynamics is a methodological approach directed toward an integrative understanding of behavior whereby the human being is viewed as an open system composed of interrelated time-varying components and information feedbacks.

THE HUMAN DYNAMICS APPROACH

The Human Dynamics approach may be broken up in four main stages. The first stage is the determination of the relevant elements associated with the human system. For instance, motivation, satisfaction, learning and performance are large, complex and coherent enough to have been considered as subsystems of the larger human system. In addition to these four subsystems, a decision subsystem has also been included. This subsystem represents the controlling agent of the individual's performance.

The second stage of the approach deals with the construction of the human behavior model. A set of propositions relative to each subsystem, backed by existing theories and empirical evidence, have been formulated.

The third step is the translation of the propositional model into mathematical equations. This step is required to facilitate the operationalization of the model and prepare the next and fourth phase: the dynamization.

The dynamic formulation, which is used for the purpose of simulation, is derived from the static mathematical model. As Coleman [5, p. 69] says for justifying the use of simulation: "the aim is to program into the computer certain theoretical processes and then to see what kind of a behavior system they generate". And he adds: "the aim is to put together certain processes at the individual and interpersonal level and then see what consequences they have at the level of the larger system".

Vol. 5
1977

THE HUMAN DYNAMICS SYSTEM

The main theoretical foundation of the model is expectancy theory [3, 7, 10–16, 18, 19, 22, 23, 28, 33, 34, 35]. The model has some similarities with those developed by Dunnette *et al.* [6], Porter and Lawler [27] and Vroom [33] but differs from them to a large extent. For instance, the model has been formulated under a set of mathematical expressions which have been translated into a dynamic format to allow simulation. Also, either new or previously studied relationships have been determined or re-formulated. In the same way, some causal relationships have been theoretically conceptualized and integrated in the system.

The model is composed of five interrelated subsystems (Fig. 1): (1) performance subsystem; (2) motivation subsystem; (3) satisfaction subsystem; (4) learning subsystem; and (5) decision subsystem. The first four subsystems are the main determinants of the human behavior's model. The fifth one relates the first four subsystems to their environment, i.e. and as an example, the control process to which the individual is subject to in his work.

A sequence of subsystems interactions can be imagined as follows: 1, The individual is motivated (subsystem b); 2, Motivation leads the individual to

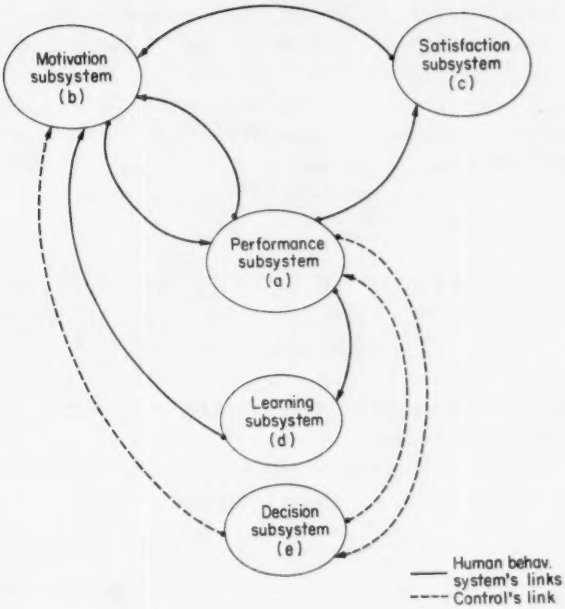


FIG. 1. *Human dynamics system.*

perform (subsystem a). Two main interactions arise: one with the satisfaction subsystem, the other with the learning subsystem; 3, With respect to the satisfaction subsystem, performance may lead to different rewards depending on the level of performance and other stochastic elements. As a consequence, the individual satisfaction (subsystem c) may vary. Performance may also influence the level of aspiration of the individual and indirectly acts upon his satisfaction; 4, With respect to the learning subsystem we start from the consideration that performance may or may not lead to a reward. From the experience of having received a reward or not, the individual is going to reevaluate his chances of being rewarded next period. This is going to influence his motivation. A same type of learning process (subsystem d) is also carried on for the reevaluation of the chances concerning the relationship between motivation (or effort) and performance; 5, Satisfaction is also going to influence the individual motivation by changing the values associated with the set of rewards. If the rewards values (or utilities) change, the individual motivation changes, which leads to a change in performance.

With respect to the decision subsystem, several other interactions take place. For example, the individual performance is compared with a given goal. If there is a discrepancy, a corrective action is implemented. For example, the amount of reward may be changed such that the existing difference between

results and goal disappears. A change of the amount of reward is going to influence motivation which is going to lead to a different performance and consequently to a new corrective action.

The whole sequence is now completed. We can see how the whole set of subsystems interacts and influences the state of the global system overtime. Before looking at the system behavior, let us turn to the study of one subsystem, that we are going to use as an illustration of the Human Dynamics approach.

AN ILLUSTRATION OF THE HUMAN DYNAMICS APPROACH: THE MODELLING OF THE PERFORMANCE SUBSYSTEM

Each subsystem has been formulated as a set of propositions. For instance, for the performance subsystem.

Proposition 1. The individual performance level is a function of his motivation to perform (or effort), his ability, and his role perception.

$$\text{Performance level} = f(\text{motivation, ability, role perception}). \qquad (1.1)$$

Motivation is defined as the resultant tendency to perform. Ability is the individual's skill. Role perception is his perception of "how to do it".

Such relationships have been proposed by Lawler [18], Porter and Lawler [27] and to a lesser degree by Vroom [33] who does not however consider role perception as being an intervening variable.

Proposition 2. Given some value of ability and role perception greater than zero, the level of performance is an inverted U shape function of the level of motivation (or effort).

The performance is low at a low level of motivation, reaches a maximum for a moderate level of motivation, and decreases for a higher level of motivation.

$$\text{Performance level} = \text{Motivation} \times (M^* - \text{Motivation}) \times g_1(\text{ability, role perception}), \qquad (1.2)$$

M^*: maximum motivation level.

This functional relationship is supported by the work of Vroom [33], where he quotes several supporting evidences [4, 25, 36].

Proposition 3. Given some level of motivation to perform and some value of role perception greater than zero, the level of performance is a linear monotonic increasing function of the individual ability.

If ability increases, performance increases too, in a proportional fashion.

$$\text{Performance level} = \text{ability} \times g_2(\text{role perception, motivation to perform}). \qquad (1.3)$$

Vol. 5
1977

The proposition is supported by the works of French [9], Lawler [18], Locke [20], Maier [21], Porter and Lawler [27] and Vroom [33] in which they argue that motivation and ability combine multiplicatively to create performance.

Further propositions are also possible but, for conciseness sake, they will not be presented here.

The equations of the Human System contain many parameters. Due to the lack of data and relevant information concerning some of these parameters, there exists no easy way to estimate their value. To solve this problem, first, the parameters have been subjectively assessed within a reasonable range of alternatives. Then, a sensitivity test of parameters change, using a factorial experiment design and analysis of variance, has been undertaken. Finally, the parameters which have been found to be system sensitive, have been integrated to the model status variables and dichotomized into low and high values to stress extreme behaviors.

The block building approach has been used through the system construction and description. This modular type of representation has facilitated model changes and improvements during and after the research. The global system has also been easily represented by the between components and between subsystems relationships.

The dynamic mathematical model, derived from the subsystems propositions has been expressed in the Dynamo simulation language [8].

For example, concerning the performance subsystem, the performance level (PEL) is dependent upon the individual effort (EFC), his ability (ABI), his role perception (ROL), and his input level (IIL). Dynamically, the relation may be expressed as follows:

$$\text{APEL.K} = \text{WEFC.K}*\text{WABI}*\text{WROL}*\text{WIIL.K}, \qquad (1.1A)$$

APEL: auxiliary performance level (no dimension);
WEFC: weighted effort component of performance (no dimension);
WABI: weighted individual ability (no dimension);
WROL: weighted individual role perception (no dimension);
WIIL: weighted individual input level (no dimension).

The rate of change in performance is going to vary depending on the value of this auxiliary equation. For example, if APEL is high, the rate of increase will also be high.

$$\text{PER.KL} = \text{APEL.K}, \qquad (1.2A)$$

PER: performance rate (1/time unit).

Knowing the rate of change in performance it is easy to derive the present level of performance index: PEL.

$$\text{PEL.K} = \text{PEL.J} + (\text{DT})\,(\text{PER.JK} - \text{PPR.JK}), \qquad (1.3A)$$

PEL: performance level (no dimension);
PPR: past performance rate (1/time unit);
DT: computational interval (time unit).

The whole system formulation expressed as a set of Dynamo equations may be found in Thiétart [31, 32]. The Dynamo formulation has unabled the model simulation and consequently the study of human behavior. However, before turning to the model simulation one final step has been mandatory, that is to say, the validation.

A validation procedure has been undertaken by using the standard experimental strategy of comparing the model's results with those of the experimental literature. This validation procedure has already been used by authors such as Abelson and Carrol [2], Hunt and Hovland [17] and Simon and Feigenbaum [29]. The experiment has consisted of test runs of the simulation model to see if the results conform to some of the assumptions used in designing the system. Then, the impact of a change of some of the system's assumptions upon the model behavior has been analyzed. In fact, it would have been undesirable to have the model output merely reflect its assumptions. A good model should have indeed revealed intrinsic characteristics of its own, independently of external assumptions which could have been introduced in the building process. It was hoped that the combining of the variables in a way conceptualized by the system may reveal provocative insights about the consolidated impact of all the variables. It was expected that the variables consolidation, more than the variables themselves, leads to specific system behavior. The final results have tended to show that all the 10 selected hypotheses were verified by the model simulation, and that the system behavior was not influenced too much by an assumption change. The validation outcome has consequently reinforced our confidence in the model behavior.

Vol. 5
1977

The description of the Human Dynamics approach is now completed. Let us turn to some examples to illustrate the contribution of such an approach for studying human beings at a global and dynamic level.

BEHAVIOR AND CONTROL: AN EXAMPLE OF THE HUMAN DYNAMICS APPROACH

"The primary aim of control is to assure that the results of operations conform as closely as possible to established goals" (Newman *et al.*, [24, p. 581]). In an industrial setting, that we are going to take as an example, one type of decision which might have a sensitive impact on the individual's behavior is the one associated with the control process. Control process has to be understood in its broader sense, that is to say: the verification of goal-performance congruence and the implementation of a corrective action, if it is required, to get the performance in line with the goal.

The effectiveness of a control process depends on some of its characteristics. For instance and as Starr [30, p. 447] points out, "the design of the control system requires awareness of how long a period can be permitted to elapse between the occurrence of the disturbance of the system and its correction".

In fact, a good control should report deviations as soon as they occur and the correction should be undertaken immediately. However, the information detection, e.g. the detection of the error signal and its transmission takes time. A corrective action, which is implemented after a time lag may take place when it is not needed any more and may consequently have an adverse effect on the system under control. To take another example, it has often been said that a deserved reward should follow the performance accomplishment without delay. In both cases, corrective action implementation and rewarding, the existence of delays seems to have a pervasive effect on the quality of control. However, depending on the type of control, we may think that the effects are different. Nevertheless, up to now, neither empirical observations, nor existing theories have given any insights about the influence of delays on human behavior and with respect to different types of control. The Human Dynamics approach provides new information.

For instance, a lag between performance and rewarding tends to create a lag in the individual satisfaction. The individual may feel the full impact of the reward only after a while, when the conditions may have changed and thus may tend to be "over motivated" or "under motivated" and have undesirable performances. In the same way, and as it can be observed on Fig. 2, a lag

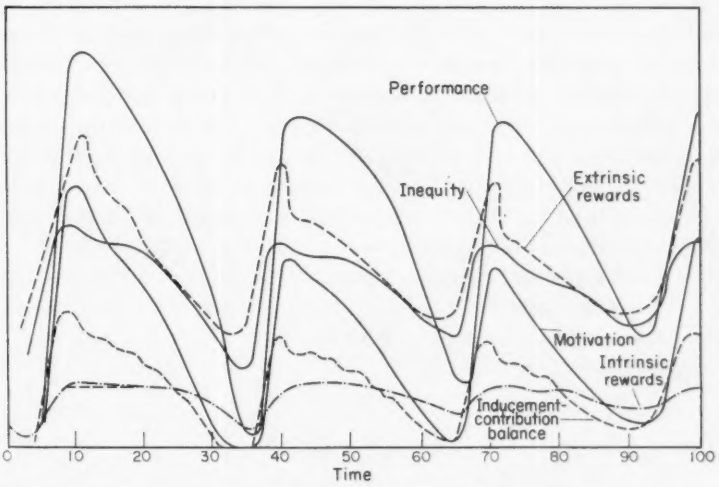


FIG. 2. *Performance and related variables. Case of a continuous corrective action with an implementation delay.*

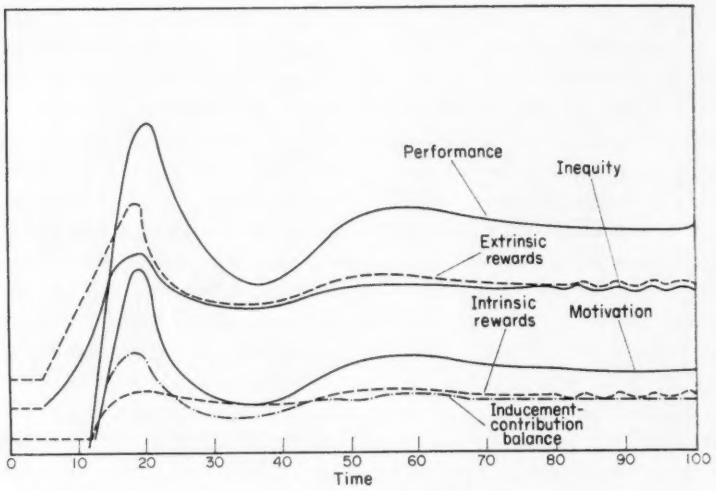


FIG. 3. *Performance and related variables. Case of a periodic corrective action with an implementation delay.*

between the instant the need for a corrective action is detected and its implementation, tends to create an oscillatory behavior. In fact, the delayed corrective action may be implemented when it is not needed any more. The performance change may be overemphasized rather than being channeled toward a same goal.

However, the impact of a delay may be different depending on the type of control. For instance, on Fig. 3, human behavior, subject to a control process relying upon a periodic corrective action, can be observed. A periodic corrective action is a periodic sequence of error detection and correction. The performance is corrected at regular time intervals. The correction, being based upon several observations will tend to generate a lesser instability than in the case of a continuous corrective action (Fig. 2). A continuous sequence of error perception and reaction, being based on a few information, will tend to associate an important weight to the last observation and will consequently be very much sensitive to changes in external conditions. The difference between the two types of correction, and the behaviors which are described, demonstrate the advantage of using such an approach for studying human beings. Several other examples may be found in Thiétart [31, 32].

CONCLUSION

By approaching human behavior at a high level of complexity, by considering the human system as being composed of interacting and time-varying elements,

by relying on simulation for dealing with the difficult problems of analysis and inference, Human Dynamics as it has been shown in the preceding examples provides new insights into the study of man.

The Human Dynamics model, that has been developed, tends to show the existence of characteristics which depend on the structure of the system itself. In fact, the system itself, rather than its components, generates new features, like the behavior instability which would have remained unrevealed if the human system had been approached in a more simplistic fashion.

Other situations and other variables could have also been investigated. The individual's performance when subject to a control process, in fact, has only been one example among other alternatives. The Human Dynamics approach is more a way of looking at problems related to human behavior and a general method of dealing with them than a rigid structure where strict procedures and well-defined steps prevail. A Human Dynamics model, relying on a block-building type of construction, may be easily modified, improved, and corrected. The theory developed along with the model may then be analyzed and tested in order to derive a better qualitative representation of the phenomenon. The Human Dynamics approach forces one to look at problems on a broader scale and in a larger perspective. The questions concerning interaction between the system's components and the inclusion or exclusion of a given element are constantly raised. The use of computer simulation obliges the investigator to define and explicate the problem he is dealing with and at the same time to pinpoint those areas where further investigation is crucial because of the lack of relevant information. Human Dynamics by relying on simulation, helps to develop theories whose descriptive and predictive powers, when validated, are enhanced in comparison to more traditional approaches. Simulation provides a mean for replicating experiments, for manipulating variables and for changing experimental conditions that no real experiment can offer. It also allows for the study of complex situations which generally predominate behavioral processes and prevents distortions in the phenomenon under analysis by avoiding too highly controlled real-life type of experiments. A greater attention to the model's formulation and a higher flexibility of the experimental design are the two most important contributions of simulation to the Human Dynamics approach.

Imperfections still persist and numerous areas for further investigations remain. Ways to refine the model itself and to extend its field of application should be found. Better support for some of its fundamental propositions are needed. Changes in the mathematical formulation of its conceptual structure may be necessary. Questions related to model's validity also deserve attention. The study of foreign structures introduced when writing the simulation program presents also some interest. Other extensive investigations should be made to use the Human Dynamics approach for studying the individual's behavior in his non task related activities.

The Human Dynamics approach aims to be a general way of looking at and

solving problems associated with human behavior. Much research still needs to be made. It is hoped that the present research will contribute to stimulate investigation in the behavioral and managerial fields and will emerge as a step for integrating the social and management sciences into one unique science of man.

REFERENCES

1. ABELSON RP (1968) Simulation of social behavior. In *The Handbook of Social Psychology* 2nd Ed. (Eds. LINDZEY G and ARONSON E) Vol. II, Addison-Wesley, Reading, MA.
2. ABELSON RP & JD CARROLL (1965) Computer simulation of individual relief systems. *Am. Behav. Scientist* **8**, 24–30.
3. ATKINSON JW (1958) *Motives in Fantasy, Action and Society*. Van Nostrand, New York.
4. BIRCH HG (1945) The role of motivational factors in insighful problem solving. *J. comp. Psychol.* **38**, 293–317.
5. COLEMAN JS (1962) Analysis of social structures and simulation of social processes with electronic computers, in *Simulation in Social Science* (Ed. GUETZKOW H) Prentice Hall, Englewood Cliffs, NJ.
6. DUNNETTE MD, CAMPBELL JP & HAKEL MD (1967) Factors contributing to job satisfaction and job dissatisfaction in six occupational groups. *Org. Behav. Human Perf.* **2**, 143–174.
7. EDWARDS W (1954) The theory of decision making. *Psychol. Bull.* **5**, 380–417.
8. FORRESTER JW (1966) *Principles of System*. M.I.T. Press, Cambridge, MA.
9. FRENCH EG (1957) Effects of interaction of achievement, motivation and intelligence on problem-solving success. *Am. Psychol.* **12**, 399–400.
10. GALBRAITH JR & CUMMINGS LL (1967) An empirical investigation of the motivational determinants of task performance: interactive effects between instrumentality—valence and motivation—ability. *Org. Behav. Human Perf.* **2**, 227–257.
11. GAVIN JF (1970) Ability, effort and role perception as antecendents of job performance. *Expl. Publication System* **5**(90A).
12. GEORGOPOULOS BS, MAHONEY SM & JONES MW (1957) A path goal approach to productivity. *J. Appl. Psychol.* **4**, 345–353.
13. GOODMAN PS, ROSE JH & FURCON JE (1970) Comparison of motivational antecedents of the work performance of scientists and engineers. *J. Appl. Psychol.* **14**, 491–495.
14. GRAEN G (1969) Instrumentality theory of work motivation: some experimental results and suggested modifications. *J. Appl. Psychol. Monogr.* **53**(2), 1–25.
15. HACKMAN JR & PORTER LW (1968) Expectancy theory predictions of work effectiveness. *Org. Behav. Hum. Perf.* **3**, 417–426.
16. HOUSE RJ & WAHBA MA (1973) Expectancy theory as a predictor of job performance satisfaction and motivation: an integrative model and a review of the literature. Working Paper, Baruch College, The City University of New York.
17. HUNT EB & HOVLAND CI (1961) Programming a model of human concept formation. *Proc. Western Joint Computer Conf.* **19**, 145–155.
18. LAWLER EE (1971) *Pay and Organizational Effectiveness: A Psychological Perspective*. John Wiley, New York.
19. LAWLER EE & PORTER LW (1967) Antecedent attitudes of effective managerial performance. *Org. Behav. Hum. Perf.* **2**, 122–142.
20. LOCKE EA (1965) The interaction of ability and motivation in performance. *Percept. Motor Skills* **21**, 719–725.
21. MAIER NRF (1955) *Psychology in Industry*, 2nd Ed. Houghton Mifflin.
22. MITCHELL TR (1974) Expectancy models of job satisfaction, occupational preference and effort: a theoretical, methodological, and empirical appraisal. *Psychol. Bull.* **81**, 1053–1077.

Vol. 5
1977

23. MITCHELL TR & ALBRIGHT D (1972) Expectancy theory prediction of job satisfaction, job effort, job performance and retention of naval aviation officers. *Org. Behav. Hum. Perf.* **8,** 1–20.
24. NEWMAN WH, SUMMER CE & WARREN EK (1972) *The Process of Management.* Prentice Hall, Englewood Cliffs, NJ.
25. PATRICK JR (1934) Studies in rational behavior and emotional excitement: the effect of emotional excitement on rational behavior in human subjects. *J. comp. Psychol.* **18,** 153–195.
26. PEAK H (1955) Attitude and motivation, in *Nebraska Symposium on Motivation* (Ed. JONES MR). Lincoln University of Nebraska Press.
27. PORTER LW & LAWLER EE (1968) *Managerial Attitudes and Performances.* Dorsey, New York.
28. ROTTER JB (1954) *Social Learning and Clinical Psychology.* Prentice Hall, New York.
29. SIMON HA & FEIGENBAUM EA (1964) An information-processing theory of some effects of similarity, familiarization, and meaningfulness in verbal learning. *J. verbal Learning verbal Behav.*
30. STARR MK (1971) *Management: A Modern Approach.* Harcourt, Brace, Jovanovich, New York.
31. THIETART RA (1974) Human dynamics. A methodological contribution to management and behavioral sciences. Unpublished Ph.D. dissertation, Columbia University.
32. THIETART RA (1977) *La Dynamique de l'Homme au Travail. Une Nouvelle Approche par l'Analyse de Systèmes.* Les Editions d'Organisation, Paris.
33. VROOM VN (1964) *Work and Motivation.* John Wiley, New York.
34. WAHBA MA & HOUSE RJ (1974) Expectancy theory in work and motivation: some logical and methodological issues. *Human Relations* **27,** 121–147.
35. WOFFORD JC (1971) The motivational basis of job satisfaction and job performance. *Personnel Psychol.* **24,** 501–519.
36. YERKES RM & DODSON JD (1908) The relation of strength of stimulus to rapidity of the habit formation. *J. comp. Psychol.* **18,** 459–482.

ADDRESS FOR CORRESPONDENCE: *Dr RA Thiétart, Associate Professor, Ecole Supérieure des Sciences Economiques et Commerciales, BP* 105, *Cergy* 95001, *France.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

Interactive LP Models in Management —An Analogue Approach

JHD WALTON

(*Received May* 1976; *in revised form October* 1976)

The history is traced over 10 years of a project directed towards extending the use of Linear Programme methods by managers and planners themselves. Both computational aspects as they affect the end user, and possible alternatives to digital computation, are discussed in the light of practical experience and of recently proposed criteria for mathematical modelling in a managerial environment. The operating principles of an LP-based electronic analogue system are described, along with some of its observed implications in terms of human behaviour. It is concluded that LP and related techniques have hardly begun to exploit their potential, both as a control medium and as a language of managerial communication; that the concept of direct optimisation is largely irrelevant and has formed a barrier against implementation; and that the field of application for LP models is far wider than can be adduced by rigorous analysis, or by working to the practical bounds of digital computation.

LINEAR PROGRAMMING—THE STARTING POINT

SOME 10 years ago, in the course of OR project work which followed an earlier career in Electrical Engineering, it occurred to the author to summarise two separate aspects of LP which were not at that time in the forefront of attention:

> "The use of a simple LP model or format often greatly improves the power to visualise and control a situation. Such models should therefore be amongst the basic tools of management".
>
> "A LP model can be translated directly into a physical analogue, which in effect computes itself, by using dials and switches linked into an electrical circuit".

These statements can be confirmed by practical demonstration, and form the basis for believing that LP is a far more powerful and universal tool than are the involved calculations and digital computer routines which in most cases represent LP in the literature—upon assumptions regarding the nature of managerial decision-making which of late have been increasingly called in question [2, 3, 5, 9, 11]. This paper describes how an experimental investigation arose first out of asking a simple, basic question: "Why are small LP models not extensively used by companies to monitor profitability, and their internal allocation of scarce resources, as the textbooks suggest?". To this question nothing in the literature appears even now to give a fully convincing answer.

In the event that the companies themselves were small, the view obtained upon questioning professional colleagues, and others with relevant experience, was that small companies simply could not afford to buy the necessary calibre of expertise for a sufficient time. Yet with increasing numbers of user-oriented LP packages being offered by the Bureaux, LP still did not seem to gain immediate and widespread acceptance as a management tool in the small company area.

Again one asked, why was it that companies with both the necessary expertise, and the volume of data handling to justify comprehensive computer facilities, did not appear to use small LP models of short term factors, for guidance if conditions suddenly changed? This question appeared to have some force, in that personal experience of some large LP type product mix models had suggested that they might be up against fundamental limiting factors of a kind familiar to the communication engineer: namely a necessary trade-off between the time-span of control and the amount of detail in the model.

One hypothesis which fitted the facts was that managers failed to be attracted to LP as a control medium because they felt it would not be a process which they could follow through completely themselves. Hence they would presumably fear loss of information at the interface between themselves and the computer in both directions of information flow: that both flexibility and confidentiality would be impaired, in comparison with existing procedures. It had already been observed that the optimising routines, associated with LP, themselves aroused some scepticism amongst managers, due in some measure to lack of understanding.

Vol. 5
1977

INITIAL EXPERIMENTS

The intriguing nature of these interface problems, if they existed, justified taking up the opportunity to gain direct experimental evidence. The starting point, in 1966, was participation in the running of a 'do-it-yourself' LP course for managers. On this course, each delegate was invited to learn the fundamentals of LP, and also to acquire a necessary facility in related areas such as multiple regression and the assessment of sampling error. Delegates then had to formulate and solve an LP model which was of some real significance in terms of their own company operations. Thereafter, when the model had been 'debugged' and initial solutions explored in discussion, they would be able to go round again in order to relate more closely to practical considerations. Of the six delegates all except one, who arrived late, succeeded in formulating an LP model and in obtaining solutions which, on the face of it, were worth £100,000 per annum or more to their companies. Yet strangely, not even the considerable publicity which attended this result resulted in a flood of new applications for the course. The inferred psychological barriers evidently did exist.

The second experiment occurred almost fortuitously in the course of a marketing consultancy exercise in 1966. At an early stage in this project it had become clear that the client might well go out of business before any marketing advice could be implemented. The situation was that following a history of profit, the client had achieved simultaneously an overall trading loss, and a state of deadlock with the local Trades Unions on account of the low level of wages being paid. Attempts had already been made to apply marginal costing, which clearly showed that some products were much more profitable than others. Since however all products showed at least some contribution to overhead, and the least profitable were in fact those by which the company was best known in the trade, the client had not been able to see product rationalisation as an immediate way out of current difficulties. The structure of the company and its products lent itself quite readily to the construction of a small LP-type Product Mix model with only 10 variables and 5 constraints, in addition to individual bounds on the variables. This model yielded a wealth of information; it powerfully reinforced the previous conclusions, based upon marginal costing, that certain products ought to be dropped; however it also showed that simultaneous adjustments to product pricing, along with marginal improvements to 'productivity' factors in the model, could offer the prospect of an almost equivalent increase in profit. This would be achieved without inflicting violent change upon either the work force or the company's sales image. The nature of these combined changes did not however figure directly in any computer print-out, or even come in a flash of inspiration to any one individual. It was rather that collective insight resulted from teaching the language of LP to a number of individual managers who despite having little confidence in 'optimality' as such, thereby came to appreciate the nature and meaning of the 'opportunity costs' implied by the computer solution. The Sales Manager for example realised that though on the face of it he had no room for manoeuvre in that he could not withdraw unprofitable items from the catalogue, he could at least decline to take the initiative in quoting for the worst individual items within the unprofitable class. Instead he could concentrate his quotations and selling effort where the greatest shadow costs appeared on the sales constraints. Also he took steps to adjust the discount structure, to ensure that a similar approach would be adopted by sales agents.

The practical effect was dramatic. Although the company at the formal level had withdrawn no individual products, nevertheless over a period of some four months the rate of sales overall increased by 27% despite a falling market. This was sufficient to avert the immediate danger. In fact, the dynamic behaviour of the model had proved to be as informative as any individual solution. The tendency of one product group to oscillate in and out of the optimum mix was eventually attributed to poor delivery performance by a component manufacturer, which occasioned recurrent spasms of wrongly sequenced and hence unproductive work. The full effect upon profit potential had not been obvious

to the management before, and as the product concerned was a new development upon which considerable hopes were pinned for the future, this result alone was of some consequence. This second experiment thus amply confirmed expectations as to the potential of the small LP model as a control medium, if correctly used, in that the result achieved had not been obtainable from conventional costing. This was presumably because costing as such does not embody directly any concept of limited resources, and marginal costing only deals with one limiting resource at a time—perhaps not always the right one [7].

Much light was shed upon the nature of the interaction between management, and a LP model employed as a control medium as described. The benefit had clearly come not from a single 'snapshot' model, still less from a single 'optimal solution', but rather from subjecting repeated solutions of progressively updated models to what was in effect a type of exponential smoothing. Managers were not expected to take immediate, drastic action on the strength of one computer print-out, but rather to make a limited, progressive response as circumstances allowed. Another aspect was the great psychological value of the routinely up-dated solution. If the indication of which way to move is obtained consistently from a number of independent sampling of company data, people will tend to judge that the indication is correct. In contrast a single computation from a larger data sample, in theory giving better accuracy, will be less readily accepted since it will not allow intuitive assessment of the degree of randomness, nor of rates of change. The practical advantage of a small model also showed clearly, it could be readily understood and the detail remembered, and in this case the volume of data was small enough to be handled by one part time clerk with reduced risk of punching errors.

Vol. 5
1977

The experience of this project thus reinforced the concept of an evolutionary approach to modelling. Since even to reconcile individual product volumes with accounting figures, in order to form meaningful moving averages, could be seen to involve lengthy discussions and adjustments it was clearly impossible to maintain at any stage that the model was perfectly accurate, or that it would lead to some uniquely valid 'solution'. Rather, one had learned to stress that a model used for control purposes constitutes not a record of the past, but a forecast for a limited period in the future. An essential need was for the client to conduct some form of sensitivity analysis so as to face up to the consequences of likely inaccuracy, and as a further precaution the accuracy and stability of important parameters should be continually monitored by graphical means.

Quite apart from obvious questions of sampling error and computational accuracy, it was only when confronted with the computed consequences of a given set of data that the client could call to mind hidden inaccuracies and defects in classification, and perhaps recognise his own tendency towards wishful thinking. To present a so-called optimal solution before the outcome of several simpler computations had been presented, and argued out, was clearly to be

regarded as counter-productive. In practice, before any computation could be done at all, the model had to look right. If one matrix coefficient for a product appeared to be twice the magnitude of another for a different product, individual managers who knew about the manufacture had to accept that this was reasonably near to the truth. To allow such doubts to persist, or to remain unrecognised, would undermine confidence and allow results to be easily discredited by those who chose to be identified with vested interests. Also, when present sales levels were substituted into the model, the derived overall utilisation of critical resources had both to appear reasonable, and to correlate with the observable queueing times which determined the company's delivery performance.

A further technique used at the time, and the importance of which has since been established, was to explore with the client the idea of partitioned single-variable solutions before proceeding further. These can be presented in a form such as: "What would happen if you temporarily dropped everything else and concentrated upon Product A?". In a small model at least, this question usually has a measure of practical meaning; it checks assumptions of linearity, and forms another safeguard against hitherto undetected inaccuracies in the matrix. This procedure, upon which has since been based the concept of the "Normalised Model", has been considered the key to success in the analogue modelling of company situations.

COMPANY MODELS—THE ANALOGUE APPROACH

The third experiment, which overlapped the previous work, was aimed at discovering whether a company model of the conventional product mix type could be formulated and effectively used without the aid of digital computation at all. This was because time delay in batch processing, and some distaste on the part of managers for involvement even in simple programming routines, had already been seen as recognisable deterrents.

Analogue methods had already been evolved and become established in feed mix computation; but the case was not entirely parallel because analogues had tended to be presented actually as computers—delivering hard copy from an input of well established and reliable data.

Consideration of other reported applications [1, 6] of analogue to the implementation of OR-based techniques suggested that the advantages of analogue would be:

1. The very fast response of the analogue, occurring within seconds, and its consequent ability to provide 'feel' as factors are changed at will, would be of great advantage. Also, it would be easier to provide for optional or 'soft' constraints [5].

2. The small size of model often used, and the amount of uncertainty to be expected in the data, meant that the size restriction and limited accuracy of an analogue would not be serious drawbacks.
3. Analogue would lend itself readily to what may be described as the 'Sub-LP' stage of data validation already described. The apparently simple, but time-consuming, check calculations carried out in collaboration with the client are not greatly assisted by conventional computer programmes.
4. On an analogue suitably constructed, the solution to an LP can be detached from the constraints and controlled independently. In other words, it is possible to depart from a basic solution and introduce a finite amount of slack on all the constraints. Though a non-basic solution cannot be optimal in normal LP terms, it is more likely than a basic solution to be implemented since it will comprise the 'room for manoeuvre' which in management models is increasingly recognised to be essential [3], and also perhaps avoid the complete elimination of variables due to a marginal cost difference.

Vol. 5
1977

It now remained to deal with the other important disadvantage of analogue: that it can handle only a very restricted range of numerical value on account of the physical limitations of the components used. This does not matter so much in, say, a feed mix application, when both the variables and the matrix terms themselves are in the form of percentages. In a company model however the variables may range in very different orders of magnitude, besides having to be in appropriate units, and the same applies to the terms of the matrix. Hence it seemed at first sight that an analogue would have to be built specially for every company situation. The question was, in effect, how to make a company model fit within a predetermined numerical form whatever the type and source of data, eliminating both the units of measurement and any numbers outside the range which the analogue is built to handle. It was decided that in view of the benefits obtainable—elimination of decimal points and of zeroes needing to be counted—the ideal arrangement would be to have all the matrix and RHS terms expressed as percentages. The variables should also be expressible as percentages, at least within their feasible ranges, or else possibly range from 0 to 1.

The method devised for achieving these numerical restrictions led directly to the concept of the Normalised model as described below. Adopting the known technique of the partitioned solution, the model is solved—normally by inspection—for each of the variables taken one at a time. Only upper limit constraints are considered. Then each column of the matrix is multiplied through by the value of its corresponding variable in the partitioned solution, and the variable itself is divided by this quantity so as to constitute a new transformed variable, which is dimensionless. The constraint rows are divided through by their RHS (constant) terms and then multiplied through by 100. This procedure

will (rigorously) leave the solutions to the model unchanged in their interpretation and (on certain non-negativity assumptions[1] which in practice are often disregarded) cause the matrix terms to range only between 0 and 100. The variables, as transformed, will on the same assumptions range only between 0 and 1 in feasible solutions. The objective function row may be normalised by dividing it through by the value it takes in the best (i.e. optimal) solution, as a convenient way of making these matrix terms percentages also, if so required; but it may be scaled in other ways, or left unchanged.

Since the scaling described, whatever the factors used, leaves the solutions to the model unaffected once they are transformed back into the old variables, the accuracy of the partitioned solution is not material to the final result. Failure of the non-negativity assumptions, or the substitution of arbitrary numbers for any or all of the values taken by the variables in the partitioned solutions, will affect only the ease of numerical handling and/or the intuitive interpretation of individual factors: the former being sometimes facilitated by deliberately rounding off the partitioned solutions to convenient integers. In general it becomes relatively easy to recognise conflicting and redundant constraints.

Once the pre-conditions had been understood and met, it became possible to construct analogue simulations of company situations such as that described. It has subsequently transpired that a simple company model can actually be formulated most easily by using the normalised form from the start. A normalised model forms of itself a readily comprehensible mathematical language in which, independently of any computational or optimising process, the limiting resources and earning power of a business can be succinctly expressed.

THE NORMALISED MODEL—ITS DERIVATION AND HANDLING

Consider a very simple Product Mix model whereby a manufacturer seeks to maximise the profit from product volumes x and y, within two process capacity constraints, as follows:

$$\begin{aligned} &\text{Maximise} && 11x + 7y = Z && \text{(Profit)} \\ &\text{Subject to:} && 7x + 6y \leq 84 && \text{(Capacity of Process 1)} \\ &&& 2x + \ y \leq 16 && \text{(Capacity of Process 2)} \end{aligned}$$

and the usual non-negativity constraints upon the variables.

[1] Strictly, the assumption is that no constraint row which is limiting in a partitioned solution shall contain a negative matrix term, since this would allow higher values of the variable to occur in solutions also containing a second variable. This is seldom troublesome in practice due to the influence of other constraints. In cases of doubt, standard digital procedures can be used to establish bounds upon the variables—particularly when similar procedures are being employed to reduce a larger model for ease of handling.

To obtain the partitioned solution for x, we place $y = 0$. Of the two constraints, clearly the second is more limiting and hence we derive the solution $x = 8$ and $Z = 88$. Setting $x = 0$ to obtain the partitioned solution for y, we obtain $y = 14$ and $Z = 98$, since in this case the first constraint is the more limiting. Applying the successive transformations described above, multiply through the first column by 8, and the second column by 14; multiply through the first constraint row by 100 and divide by 84; multiply through the second constraint row by 100 and divide by 16. Finally, define new variables X and Y such that $X = \frac{x}{8}$ and $Y = \frac{y}{14}$

The transformed model then becomes:

$$\begin{aligned} &\text{maximise} && 88X + 98Y \\ &\text{subject to} && 66\tfrac{2}{3}X + 100Y \leq 100 \\ &&& 100X + 87\tfrac{1}{2}Y \leq 100 \end{aligned}$$

The optimal solution is $X^* = 0{\cdot}3$, $Y^* = 0{\cdot}8$ and $Z^* = 104{\cdot}8$, as may be verified by the usual processes of computation.

The interpretation of this formulation is, for example:

> "There is an available 'extreme' production option to produce nothing but product 1. This will be limited to 8 units per period by the influence of capacity 2; and because of this, only two thirds at most of capacity 1 can ever be absorbed by product 1 alone, and the profit cannot exceed 88."

In the event that such a model may be used where time does not allow the computation of an optimal solution, an upper bound on the objective function is nevertheless easily calculated (by partitioning the dual). This also shows whether the best of the partitioned solutions is in fact already optimal. In summary, the dual partitioning method is as follows:

(a) Write against each matrix term the quotient obtained upon dividing it into the product of its associated RHS and objective function terms. In the numerical example the first quotient would be $\frac{88 \times 100}{66\frac{2}{3}} = 132$.

(b) Mark the highest quotient in each row. A row containing a zero in the matrix can be disregarded, for this purpose only.

(c) Select the lowest of the marked quotients. This is the required upper bound.

(d) Compare this upper bound with the lowest quotient to be found in the same column. If the two figures are the same, then the best partitioned solution is optimal and no further computation is required. In the numerical example, these figures are 112 and 98 respectively, so that neither of the single-variable partitioned solutions is optimal.

The method described above presupposes that all of the matrix terms are positive, and becomes less useful—though still valid—when there are a large number of zeroes in the matrix.

Vol. 5
1977

SHORT-CUT MODELLING

The derivation of a 'normalised model' from an LP-type Product Mix model in conventional form has been described. It is often more convenient, and much quicker, to formulate the model in normalised form in the first place. One first represents the variables in the model as representing the degree to which each of several available extreme policies is adopted. Assuming a normal type of product mix model, one asks for each product type:

> "What would happen if other products were dropped for a time, and this one product type were made to the maximum possible extent within present resources?"

By enquiring about each potential limiting resource in turn, one discovers which is the most restrictive for that variable, proceeding then as follows:

(a) Write the figure 100 in the matrix cell which relates the variable to the most restrictive limiting resource. The interpretation is that under the envisaged 'extreme' policy of producing that product to the maximum extent which existing resources will allow, 100% of that resource would be used.

(b) Write below the matrix the extreme ('BASE') value which would attach to the variable in these imagined circumstances with 100% of the most limiting resource being used, showing also the units in which this variable (i.e. product flow) is measured.

(c) Calculate the percentage of each of the other limited resources, depicted as rows of the matrix, that would be used in these same imagined circumstances, and enter the percentage figures in their respective matrix cells. All these figures must be less than 100, or the resource concerned would have been the most limiting in the first place and an error would be diagnosed.

(d) Write in for the objective function term the value of the Profit (or whatever this row depicts) when the variable takes on its 'BASE' value.

(e) Repeat for all the variables in turn.

It will be seen on reference to the numerical example that the answers to these questions will have filled in all of the model except the constraint RHS terms which are taken to equal 100, being 100% of a resource amount which has to be specified and then written in, with its units of measurement. The model can thereafter be further elaborated as follows:

(f) Extend the matrix to deal with any further constraints, which may be of lower limit type if necessary, either using a RHS term equal to 100, or perhaps duplicating an existing row with a RHS term less than 100 to constitute an 'interval constraint' [4].

(g) Add any required bounds upon the individual variables, remembering to divide the RHS term by the 'BASE' for the variable before entering, so

as to maintain a consistent scaling. (Similarly, after computing any solutions or values for the transformed variables, these values each have to be multiplied by the appropriate 'BASE' to restore the original units of measurement).

(h) Devise a scaling for the Objective Function row. It is possible to divide through the row by the optimal objective function value, as indicated above. However it is more usual to estimate in a company model the break-even point—the value of objective function which the company needs in order to survive—and divide through by that, multiplying the answers by 100 to express them as percentages. The terms in the objective function row then have physical meaning, answering the question "Could this product alone carry the company, if enough could be sold?". Such conclusions can be checked by other means to confirm the accuracy of the model.

The procedure outlined above, applied with reasonable care and common sense, will allow an approximate but usable model of the earning capability of many types of business organisation to be built up in a matter of hours. The resulting model may be applied with equal ease to visual assessment, construction of bar charts, validation of budgets and production schedules, computer optimisation, or providing input to an analogue. Since every factor becomes expressed essentially as a percentage of ultimate capability the rationale is obvious to management personnel, and indeed such a model is discovered to form an effective channel of communication between the various managerial functions. Naturally in a real situation the estimated percentages, which build up the matrix, would soon be replaced by proper analysis of records. Also, they would have to be 'validated' by substituting data for known past periods. However the point is made that this technique is so powerful that even without 'proper' data, valuable insights will be obtained in situations of economic change or enhanced competition, or where severe pressure exists [2].

Sometimes different production policies for the same product may be represented by separate variables—for example production in large or small batches. The economics of batch quantity determination may be represented as a trade-off between cash and storage capacity on the one hand, and the drain on production capacity occasioned by over-frequent machine setting on the other. The inclusion of these factors as separate resource constraints allows the use of a linear model to evaluate Economic Batch Quantity in suitable cases, in contrast to more general methods such as that of Reeves which require Convex Programming methods [8].

ANALOGUE LP—A PRACTICAL APPLICATION

The immediate sequel to the experiments described was an application in the mineral processing field, which extended over a period of some years and

Vol. 5
1977

proved to have far-reaching effects. In the south west of England there exists a considerable industry devoted to the production of china clay. This widely used commodity is sold in many grades, varying from pure and highly refined materials used in medicine or as a filler in high grade paper, to fillers and coating agents used in plastics, rubber, and agriculture—quite apart from the production of ceramics from which the name derives. Deposits available to be mined vary in their chemical composition, and there is a continuing day-to-day problem in allocating the production from many separate working faces to individual types of product clay. The allocation problem is readily formulated as a LP containing,

Quality constraints relating the properties of blended products, after processing, to the properties of individual constituents.
Constraints upon overall tonnage, expressing the fact that one cannot use more of a given feed clay than is being mined or stored.
A set of constraints to ensure that even in the short term, all clay produced must be blended into a saleable product, or else disposed of at extra cost.

The company concerned found that the LP, once formulated, could be easily solved using the in-house digital computer; but that the solutions generated did not gain ready acceptance. Accordingly, a prototype analogue model was constructed which would embody all of the quality constraints simultaneously, continuously display the relevant data and, in a matter of seconds, provide a feasible solution within the given set of constraints. The analogue machine was then placed in the hands of an Area Chemist, whose responsibility was to programme the operation of one of the Company's blending plants. It came into use immediately, and has remained so. At an early stage the statement was made that it appeared to be recovering its capital cost every seven weeks. The design and building of larger analogues for similar purposes followed in subsequent years. The main technical feature of these larger machines was their ability to formulate several grades of product clay simultaneously, using limited quantities of high grade feed clays to impart the desired properties. They also remain in use.

The benefits conferred by on-the-spot LP modelling were described by the Company as:

(a) A reduction in the amount of material which could not be absorbed into the blending process; this material would otherwise be wasted.
(b) Ability to produce to a given set of market demands with greater tolerance as to the properties of incoming material—thus making it easier to avoid mining in sensitive areas, and presumably reducing costs.
(c) More flexible output in that it became easier to maximise production of particular grades which happened to be in demand, within constraints of production capacity and available feed tonnages. This would allow better customer service from given levels of stockholding.

(d) Greater average value of output from restricted plant capacity.
(e) Less variability in the product, and hence greater ease of storage.
(f) Ability to respond rapidly to emergencies such as plant breakdown, and to human errors should they occur.

THE INTERACTION BETWEEN MANAGEMENT AND MACHINE

During the projects described a number of lessons were learned regarding the conditions for satisfactory interaction between a machine embodying a LP model on the one hand, and a system containing human managers on the other. There was the obvious desirability of a machine affording instant access, rapid response, and which would display (not merely store) all its input data. There were however deeper considerations.

Firstly, in the case of clay blending, it became clear that there would be conflict between a single objective function and other control objectives, the latter varying according to the plant conditions. This feature recalled an impression gained during the company modelling experiments that Profit was being used as a low limit constraint rather than in the maximising sense.

Vol. 5
1977

How could varying objectives be effected in the model without the use of complex weighted objective functions? The approach in practice was:

(a) First to set up all of the known, formalised constraints and see if a feasible solution could be found at all; if not, this suggested a need to revise production targets. However, before doing so, it was worth while to investigate, using the model, whether the required output could be obtained by reducing safety margins—using personal knowledge of plant performance factors coupled with increased supervision. A feasible solution found in this way, and which appeared satisfactory to the manager concerned, was likely to be adopted.
(b) In the more frequent case where a feasible solution could immediately be obtained, and there was clearly an extensive solution space within the constraints, it became possible in principle to 'optimise'. This was most likely to be achieved by imposing tightened or even completely new constraints, at the manager's discretion, until the solution space had evidently shrunk to zero and only a single solution remained feasible. Thus immediate short-term plant objectives, such as remedying current shortages, could be accomplished in addition to formalised cost reduction —such as would have resulted from a conventional objective function.
(c) At the end of the intervention, managers tended to try and ensure that constraints reflecting quality specifications were even tighter than they needed to be, to give a consistent product and added robustness against uncertainty.

It appeared that the control scheme would have had little chance of success without such flexibility. Further thought suggested that optimisation as such was really being carried on at a higher level in the control hierarchy, by those who were seeking to optimise the use of extractive and transportation capacity, and to exploit market opportunities—all of which created the shortages which Production Management were continually concerned to correct. Such correction then served to relax the 'production' constraints which bore upon the higher level optimisation process. It seemed important to understand the nature of these constraints in setting up any system at the production level, to ensure that it would respond correctly.

Next there arose a fundamental technical point, that the clay blending situation could not be accurately represented as a Linear Programme at all. On examination it proved that whenever the analogue machine was being used by an individual having management insight and responsibility a number of logical, as opposed to quantitative, constraints were being introduced by operator intervention: for example the decision to exclude a given feed clay from a particular blend despite the fact that the Linear Programme 'computation' wanted to bring it in. This was because certain types of plant restriction could not easily be reflected within a linear model. In this way a rather astonishing point was observed, that the operating personnel were obtaining by analogue solutions satisfactory to themselves (and to the plant management) to problems far more complex than had been explicitly covered in the system design. The advantage of being able to allow for short term and even personality factors also appeared. During the formulation of a clay blend, one of the operators was seen to be looking out of the window. This proved to be because he was assessing the probability of rain, which could marginally affect one of the factors in the blend but which could easily be corrected by his skilled knowledge of the process. The talent of individuals in setting up special pipeline configurations was likewise on occasions taken into account.

Finally, on a purely human note, it was found that the analogue system served to enhance feelings of personal involvement and responsibility, which are not always characteristic of centralised computer control systems.

RESULTS AND CONCLUSIONS REACHED

Bringing the projects described up to the present day, what began as a series of questions and experiments has developed into a coherent set of techniques—including a general approach to building electrical analogues with multiple negative feedback paths and a controlled dynamic response [10], and to tailoring real life data in order to fit analogue hardware of standardised type.

At the centre of it all is the idea of broadening the concept of LP to place small LP models directly in the hands of managers. At the lowest level, this

would allow the use of LP as a descriptive language, and an insight into what factors would need to figure in a LP model even if the detailed work is never done. This knowledge alone suffices to solve many problems in practice, and to explain some of the behaviour of economic systems in likely future circumstances such as shortage of raw material and energy. At the other end of the scale lies the possibility of exercising actual day-to-day control over an organisation or process, by feeding in information to constitute a LP model and noting which way the solutions move. The resulting action may be to implement such solutions, or to correct some change in the system which made the solutions happen, or whatever is most appropriate to maintaining stability and performance. Electronic analogues form an option, as against the more usual digital computation, where fairly small models are to be handled repetitively over time, and where speed and flexibility are of the essence due perhaps to changing or uncertain input data. Some analogues may have to be specially built. In other cases, such as when using the concept of the 'normalised' company model, standardised analogues are available which can even be hired. There are many applications where the capital cost, let alone the hire charge, can be recouped in a matter of weeks whether analogues or computer terminals are used.

The context may of course be actual control, using an ongoing model of system parameters, or providing to take over control only in emergency circumstances, or even merely to indicate the need for changing over from one control mode to another. For example one may envisage a small LP-type model which would summarise the principal flow demands and capacity restrictions upon a transportation network or production facility. If this over-simplified model would not at a given time generate a feasible solution within its own subset of constraints, then logically the real system with its additional constraints would be even less capable of sustaining a feasible flow pattern. Thus there would be good grounds for initiating a change in the control pattern of the real system—for example opening new flow paths or taking steps to reduce demands upon the system—until such time as a model reflecting the original control pattern had once again confirmed the existence of a finite solution space. Such a two-stage system would have the advantage of being a 'predictor', of sensing the impending overload in a model operating on forecast data in advance of real time—whereby it may be prevented from happening—instead of responding only to an overload which has already happened and where the most telling counter-measures might no longer be practicable.

It is necessarily implied that unlike the examples so far given, systems may be modelled not because there is any power or even desire to control them in detail, but simply as an aid to understanding them and predicting their response to change—whether individually or globally. Firms may well seek to model not only their internal processes, but also their customers, sales agents, suppliers, and competitors. Such models, in that they are normally built from

incomplete or even misleading data, can derive particular benefit from the highly interactive nature of a LP analogue.

CRITERIA FOR FUTURE APPLICATIONS

It is thought that hallmarks of situations where simplified Linear Programming in the hands of managers themselves, and analogue methods, could apply, are as follows. The list is by no means exhaustive.

1. Uncertainty of data in control situations where time is of the essence, and where a multiplicity of constraints exist which can at least be approximated by linear inequalities.
2. A need, within an established LP application, for exploring complex systems of contingencies having logical inter-relations which cannot be satisfactorily predetermined in detail, but which will be apparent at the time to an informed user.
3. Situations which can be structured, but not rigorously solved, by LP because they are really far too complicated—for example a blending situation with logical or integer constraints—and where the aim is a marginal improvement in human decision-making.
4. The general case where single objective functions and/or extreme point solutions are inadmissible, and 'satisficing' is the requirement.
5. 'Sub-LP' situations of load balancing, and wherever the models employed are (or can be) so simple that the cost and delay of more formalised LP would act as a deterrent.
6. Sensitive situations, requiring utmost confidentiality.
7. Situations, often arising in the marketing area, where the structure of the model itself is the prime uncertainty.

REFERENCES

1. Calvert R & Mildwater J (1965) An analogue computer for forecasting and stock control. *Electronic Engng* 660–666.
2. Chilver H (1975) Wider implications of catastrophe theory. *Nature* **254,** 381.
3. Eilon S (1971) Goals and constraints. *Jl Mgmt* **8,** 292–303.
4. Eilon S (1962) *Elements of Production Planning and Control*, Chapter 12. Macmillan, London.
5. Kendall JW (1975) Hard and soft constraints in linear programming. *Omega* **3,** 709–715.
6. Lewis CD (1965) Iterative analogue computation applied to inventory policy simulation. *Comp. Jl* **8** (2).
7. Palmer RE & Taylor AH (1969) *Financial Planning and Control*, p. 204. Pan Piper.
8. Reeves GR (1976) Inventory decisions in product mix models. *Omega* **4,** 339–359.
9. Simon LS, Lamar C & Haines GH (1976) Managers' uses of models. *Omega* **4,** 253–264.
10. Walton JHD (1971) British Patent No. 1224871 *et al.*
11. Williams HP (1975) Formulation of mathematical programming models. *Omega* **3,** 551–556.

ADDRESS FOR CORRESPONDENCE: *JHD Walton, Esq., 2, Broadfield Road, Folkestone, Kent CT20 2JT, UK.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

A Method for Short-Term Profit Analysis[1]

WH GOLDBERG
RLM DUNBAR
I VERTINSKY
International Institute of Management, West Berlin
CC HUANG
Memorial University, Newfoundland, Canada
W STANBURY
University of British Columbia, Canada
A ERICSON
University of Gothenberg, Sweden

(*Received July* 1976; *in revised form November* 1976)

The focus of this paper is a description of a methodology for identifying corporate strategies which explain short-term profit performance. The methodology is illustrated through its application to the Swedish textile and clothmaking industry. There are three important features of the methodology: (1) A conceptual framework of the firm which cuts across the traditional fields of management. (2) Concepts derived from the framework are linked to operationalized variables on the basis of their empirical patternings. (3) A quasi-experimental design is used in order to provide causal inferences concerning short-term profit performance.

INTRODUCTION

ECONOMISTS investigating the behavior of firms have focused mainly upon equilibrium phenomena and rational decision-making under perfect information. Clearly, a study of short-term behavior in the face of uncertainty and bounded rationality requires an alternative framework for analysis. Examination of the management literature reveals the existence of several relevant frameworks each focusing upon different dimensions of behavior and structure. A step in the direction of obtaining an holistic model of the firm would link these disaggregated conceptual models into one integrated framework.

[1] Authors listed in random order.

In this study, we attempt to link concepts derived from the behavioral theory of the firm with concepts developed in studies of the functional areas of business such as finance, marketing and accounting. To identify links between concepts and variables, 36 indicators of firm posture and behavior were identified. Data from a sample of 80 firms over a period of 7 years were collected and used as the empirical basis for defining indicators (variables) for the integrated conceptual framework (i.e. the set which is the union of the functional and behavioral concepts appearing in the literature).

The statistical method employed to make the link between abstract theoretical concepts and measurable, well-defined variables is factor analysis. On the basis of variable patterning, associations between meanings of concepts and measurements can be inferred. In order to derive, as far as possible, independent concepts, varimax rotation of variables was performed. Variables may assume different meanings in different time periods as a result of the changing constellations with other variables. In order to investigate the consistency of these constellations of variables over time, we have introduced the concept of stability Stability in this paper is taken to mean constancy of pattern of association between variables over time. Stable patterns are interpreted according to the meanings assigned to the factors.

Vol. 5
1977

The first step in the analysis required the identification of a cross-functional framework for studying the behavior of firms [7].

Cyert and March [3] focused on short-term behavior of firms. They argued that as firms adapt only slowly in the short-term, particularly as far as external demands are concerned, there is ordinarily a disparity between the resources available to them and the payments they must make to the external environment. They called this surplus of resources 'organizational slack'. Clearly, organizational slack can take many forms. Cyert and March saw the concept as a general but useful hypothetical construct which could help explain the adjustment processes firms use to cope with changes in the environment such as movements in demand. They hypothesized that firms absorbed excess resources in relatively good times thus reducing profits. During relatively bad times, the accumulated slack provided a pool of resources to meet ongoing demands on organizational resources and to deal with emergencies

Generally speaking, organizational slack may be thought of as a complement to the equally vaguely defined concept 'organizational efficiency'. While Cyert and March hypothesized that in practice and over time, organizations accumulate organizational slack, many social–psychological discussions argue implicitly or explicitly that they should seek to improve efficiency [6], and thus, presumably reduce organizational slack. As this is done, it is also usually implicitly assumed that stated profits will increase.

Numerous aspects of organizational structure have received attention in the behavioral literature. For example, there are many studies on the effects of organizational size in the social–psychological literature [8]. For business organ-

izations, it has been found that as size increases, so does functional specialization and a tendency to adopt more standardized procedures. Intuitively, one might expect greater profits as a result of the economies of scale made possible by this increased rationalization of the work process through specialization and standardization. In contrast to this argument, some theorists have pointed out the tendency of large organizations to become sluggish in their responses to rapidly changing environments. There is little empirical evidence which unequivocably supports either of these hypotheses.

These theories, while developing a rigorous theoretical framework have rarely been empirically tested, i.e. their concepts linked to measurable variables. In contrast, work in the fields of marketing, finance, and accounting has focused upon development of concepts based on the empirical patterning of variables. Examples include the linking of financial ratios to form the concepts of financial liquidity [1], long-run solvency, short-term capital turnover, and long-term capital turnover [5], and financial leverage [7].

Using data from the 80 Swedish firms each year of the seven year time horizon 1966–1972, factor analysis with varimax rotation was used to derive our conceptual framework. Table 1 presents a summary of the stable links found between variables and concepts. The links are obtained on the basis of the dominant interpretation assigned to the variables receiving the highest load on each factor with a threshold of 0.6. The latter threshold leads to the elimination of several variables as it was impossible to infer their meanings on the basis of this rule.

THE QUASI-EXPERIMENT

Having identified these dimensions of firm behavior which we will consider to be the independent variables in our quasi-experimental design, a new dependent dimension must be added, i.e. that of profit performance. Two profit performance indicators were considered:

1. operating profit/total assets;
2. net profit/net worth.

The first indicator was suggested by Horrigan [5]. It focuses upon operating returns on invested assets, independent of the efficiency of financial management. The second indicator takes financial management into account. It measures net equity returns. Pearson correlations between the two profit indicators were significant and positive ($r^2 = 0{\cdot}3$: $P \leq 0{\cdot}01$). As in the subsequent analysis only the relative position with respect to the average of each indicator is important, we tested group-overlaps in classification based upon each indicator. The test indicated 70% over-lap among the two classifications. This was considered a satisfactory amount of agreement, and the subsequent analysis is based upon the first indicator of profit.

TABLE 1. A LISTING OF THE EMPIRICALLY ESTABLISHED LINKS BETWEEN CONCEPTS (i.e. FACTORS) AND OPERATIONALIZED VARIABLES

Factor	Variables
1. Lack of organizational slack	Sales/fixed assets
	Sales/total assets
	Fixed assets/number of employees (loading negatively)
2. Sales efficiency	Sales/inventory
	Personnel costs/sales (loading negatively)
3. Size	Total assets
	Working capital
	Sales
	Net worth/total assets
4. Short-term liquidity	Current assets less inventory/current debt
	Current assets/current debt
	Working capital/sales
	Working capital/total assets
5. Financial leverage	Long-term debt/net worth
	Long-term debt/total assets
6. Increased market share	Percentage change in market share
	Percentage change in sales
7. Independent factors	Net worth/total debt
	Net worth/long-term debt
	Net worth/fixed assets
	Net operating profit/interest payments
	Sales/accounts receivable
	Sales/net worth
	Number of employees/sales
	Investment/total assets (1 year ago)
	Investment/total assets (2 years ago)
	Change in sales/change in costs
	Cost/sales

Vol. 5
1977

To obtain strong causal inferences, however, it is not possible to rely upon statistical association, unless a proper experimental design has been implemented. Unfortunately, this necessarily large-scale industrial experimentation is not feasible. It is, therefore, useful to approximate, as much as possible, a proper experimental design by using 'nature' (in our case, the sectoral economic environment) as the provider of 'treatments', and by using some structural properties of the population as approximations to 'controls'.

The process of identifying such an analog has been termed 'quasi-experimentation' [2]. Figure 1 describes the basic features of the quasi-experimental design used in this study. The particular features of the data which permitted some causal inferences (rather than just indications of associations) were the following:

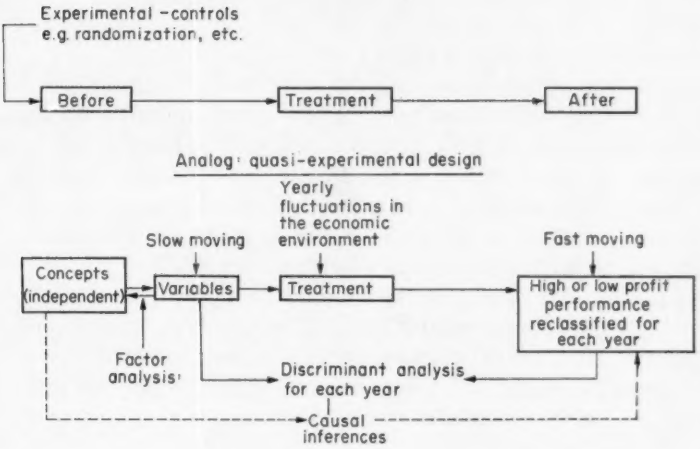


FIG. 1. *A flow chart showing the analog used in this study to create a quasi-experimental design.*

1. Dependent variables (style and posture variables) were characterized as 'slow' variables while performance variables were 'quick' variables. Relative positions of companies along dimensions of posture and style did not vary much over time, but relative profit positions did vary substantially.

2. Economic environments changed in a marked way during the horizon studied, providing the needed sequence of treatments.

These properties of the data permit the causal linking of posture and style to performance as a function of a given economic environment (the treatments).

For each year of the period of the study, firms were classified and reclassified into those with above and below average profit performance. A step-wise discriminant function analysis was then employed, with $F = 1$ as the default value to find those variables which distinguished between high and low performance. Each year was marked by differing environmental circumstances (i.e. 'treatments'). Variables with high discrimination of performance values were identified. These slow-moving variables are assumed to explain rapidly changing profit performance. In addition, and in order to obtain a basis for comparison with long-term profit behavior, averages of all independent variables were used as potential discriminators between firms with above and below profit performance over the total horizon of the investigation.

EMPIRICAL OBSERVATIONS

The environment: sectoral performance

The cloth-making and textile sector was characterized over the 7 years of our study by wide fluctuations. A surge of sales and profits in 1967 was followed by a sharp decline in sales in 1968. That year a number of firms closed down and the industry as a whole suffered a 16% drop in profits. Sales recovered in 1969, but an even greater acceleration of costs kept profits falling. This trend continued until 1972 when cost increases came under control, although sales trends then reversed. In 1972 profits increased by more than 30%.

Turning, now, to the results of the discriminant analysis, the first impression was that the variables which accounted for differences in profit performance did indeed vary as a function of changes in the environment and were often different from variables associated with long-term profitability. As perhaps would have been expected, these last included increased investment in new equipment and lower costs relative to sales.

It will be recalled that during 1966–67 the industry experienced an increase in sales. Variables associated with the factor, 'lack of organizational slack', were associated with better than average profit performance during this period. This suggests that at least during an upswing of sales, organizational slack associated with the under-utilization of organizational assets hinders rather than helps profit performance, for it indicates that not all opportunities are being exploited.

1968 was a year of sales and profit crisis for the industry. The difficulties remained to a greater or lesser extent through 1970. During this period, high short-run liquidity was associated with below-average profit. Firms did better when they emphasized sales efficiency, by reducing personnel costs relative to sales and by reducing inventory relative to sales. Similarly, those who were able to increase their market share also had above average profits. Generally, the discriminant analysis suggests that an aggressive management approach towards improved efficiency, along with the elimination of excess working capital were factors associated with improved profit performance during this difficult period.

By the end of 1971, recovery seemed to have started for the industry. Profits were able to increase substantially in 1972. One of the variables measuring managerial efficiency is associated with greater profitability, and more organizational slack is again associated with lower profits.

DISCUSSION

The discriminant analysis does suggest that at different times, and contingent upon environmental conditions, different variables are associated with different levels of profit performance. Whether management should attempt to eliminate

Vol. 5
1977

organizational slack during an upswing in sales, and in this way improve profits, seems questionable. As Donaldson [4] illustrated in a number of financial case studies, when environmental conditions turn against a firm, as they did against Swedish textile firms in 1968, this accumulated organizational slack gives the firm some leeway to manoeuvre and explore, for example, new marketing opportunities. Similarly, during a period of lower sales, higher liquidity may reduce profits but it probably enhances chances of surviving. In practice, it is necessary to take both profitability and survival criteria into account.

The method used in this study seems to offer promise for helping our understanding of firm's short-term adjustment behavior to environmental conditions. We intend to carry out further studies using similar methods but based on different data in the hope of improving our understanding of the determinants of short-term profit performance.

REFERENCES

1. ALTMAN EI (1968) Financial ratios, discriminant analysis and prediction of corporate bankruptcy. *J. Fin.* **23**, 589–609.
2. CAMPBELL DT & STANLEY JC (1966) *Experimental and Quasi-experimental Designs for Research.* Rand McNally, Chicago.
3. CYERT RM & MARCH JG (1963) *A Behavioral Theory of the Firm.* Prentice-Hall, Englewood Cliffs, NJ.
4. DONALDSON G (1969) *Strategy for Financial Mobility.* Division of Research Graduate School of Business Administration, Harvard University, Boston.
5. HORRIGAN JO (1966) The determination of long-term credit standing with financial ratios. *Empirical Research in Accounting: Selected Studies, J. Accounting Res.* **4**(suppl), 44–62.
6. KATZ D & KAHN RL (1966) *The Social Psychology of Organizations.* Wiley, New York.
7. PINCHES GE & MINGO K (1973) A multivariate analysis of industrial bond ratings. *J. Fin.* **XXVIII** (7), 1–18.
8. STARBUCK WH (ed.) (1971) *Organizational Growth and Development.* Penguin, Harmondsworth.

ADDRESS FOR CORRESPONDENCE: *Professor Walter H Goldberg, International Institute for Management, Griegstrasse 5–7, D-1000 Berlin 33, West Germany*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

Computer-Based Scheduling of Technical Resources

JO JENKINS

CDJ WATERS

Imperial College, London

M GRAHAM

M HOWELLS

BBC Television

(*Received November* 1976)

This paper considers the design of a Scheduling Sub-System which could be attached to an existing computerised data processing system in order to enhance its overall operation. It draws heavily on the experience of the authors in designing a system to assist in the scheduling of video tape machines within BBC Television.

INTRODUCTION

THE GROWTH in power of computers during past years has given an increased opportunity for organisations to benefit from the development of a completely integrated Management Information System (MIS). The concept of an integrated system does not necessarily imply that it should either be designed or installed in a single step, but it does indicate that any new components of a developing system should be designed to interlock with those parts of the system already in operation.

In most real situations, for instance, the requirements of the system either grow, or only become known, after some period of operation, while the facilities and characteristics required of the system will almost invariably alter with time. Furthermore, on a practical level, the necessary resources (capital, labour and equipment) will not often be available at any one time, but will accumulate slowly over a period. It may be implied, therefore, that while it is possible to consider MIS as a discrete entity, incapable of either flexibility or growth, it would be foolish to do so. In fact, one of the major advantages of a well designed

MIS is its ability to make allowances for changes as they arise. Because of the desirability of integrated growth there is a continuing demand for modules, or subsystems, which can be attached to an existing Management Information System to enhance its total operation.

One practical example of the type of situation that could benefit from the use of computers arises in the scheduling of video-tape machines in large television broadcasting organisations. Such a problem, arising in the British Broadcasting Corporation (BBC), has already been described in some detail [5], and it is beneficial to take this example to illustrate a practical use of the scheduling subsystem developed.

MIS WITHIN BBC TELEVISION

BBC television has as its objective the production and transmission of television programmes, on two channels, for about 20 hr a day. This involves it in the making of some 6000 complete programmes a year at a total cost well in excess of £100 million. In the mid 1960's the BBC was faced with the problem of increasing costs, caused by increased hours of transmission, the growth of BBC 2, the introduction of colour services and the general effects of inflation, while its income, controlled by Government policy, was to a large extent stationary.

Vol. 5
1977

It was obvious that some major rethinking was necessary within the organisation to overcome this trend for an increasing budget deficit. Towards this end the initial stages of a Television Management Information System (TMIS) were formulated with the explicit objectives of determining:

(a) that the service was making the best use of its revenue;
(b) whether productivity of resources could be increased;
(c) whether complexity could be reduced.

It was concluded at that time that these objectives could only be achieved with the aid of computers. It followed that in October 1969 the first computer operations were introduced into the Corporation. Not surprisingly this System had undergone continuous developments with the purpose of extending its scope and making its overall performance more sophisticated. Bearing in mind objective (b) of TMIS mentioned above, an obvious extension to the existing system is that of scheduling certain scarce technical resources, among them video-tape machines.

In particular two valid reasons can be put forward for the development of a new system for scheduling these machines. In the first place the machines are comparatively expensive both to buy and to operate. For example, in 1975, a major manufacturer of video-tape machines quoted a purchase price of up to

£75,000 and up to £50 an hour running costs. At these prices the BBC could have a capital outlay in London of about £2 million and £4 million. A small percentage saving in these costs would obviously produce a sizeable cash benefit. Secondly, the old clerical system for scheduling the machines is reaching the end of its ability to cope with the increasing demands put on it. Figure 1 illustrates the rise in the number of machines scheduled. While the system copes adequately during the summer, at times of higher demand the system begins to break down, with schedules getting poorer and outside charges to hire extra machines rising (if these extra machines are available for use).

The scheduling procedure starts some 6–8 weeks before the video-tape machines (VTs) are needed. An 'allocation clerk' enters dates, page and machine numbers on a standard form used to produce the schedules. Details of regular or 'block' bookings are then added in pencil, these being read from a 'block bookings file' which is kept up to date by frequent checks. Also at this time machine maintenance times, regular playback sessions, and advance notice of special facilities needed for major events are added. Finally 'transmission plans' and 'current studio arrangements' are used to fill in any further machine requirements known at this stage. Five weeks before the machines are to be used, a 'recording assistant' takes a week's draft schedules from the allocation clerk and completes the scheduling during the following 5 weeks. Request forms for machine time from the production departments are examined, and any discrepancies or problems are cleared up over the telephone. At this point new recordings are given a unique identifying number. From transmission and studio usage already scheduled, the staffing pattern for machine operators is worked out. Spare time for machines and men is then filled as efficiently as possible with the remaining requests for machine usage. In order to do this, and still keep within the constraints of the operation, it is necessary to repeatedly change and rearrange the bookings. This is particularly so as late requests are continually arriving, and each new arrival may necessitate many changes in the

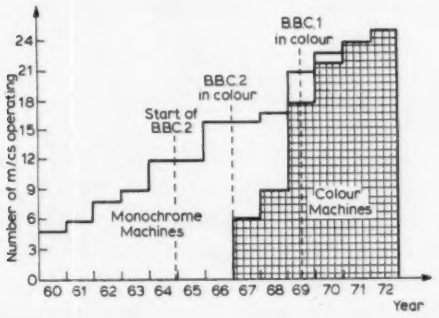


FIG. 1. *Showing the number of video-tape m/cs in operation*

schedules. In practice these alterations are made by continually rubbing out and rewriting, in pencil, the schedules already produced.

Rather than build-up a backlog of requests, excess work may be put out to external hire. Rearrangements are then necessary to the schedules to determine the cheapest means of processing the extra work, bearing in mind that hire charges are more expensive than equivalent internal operations.

Although there is a general order of priority for machine use, with some functions (notably transmissions) being considered more important than others, this may be overridden at any time by unusual circumstances. One of the main characteristics of this particular scheduling problem is that the situation is extremely flexible, with hard and fast rules being difficult to determine. The subsystem must be able to cope with this flexibility, and the schedulers must maintain complete control over the operations in order to allow for the subjective decisions necessary.

The scheduling process is complicated by the fact that more information and requests are arriving by telephone and post all the time. This includes the updating of transmission plans five times during the scheduling (at 8–9 weeks before transmission date, 5–6 weeks, 3–4 weeks, 2–3 weeks and 10 days). The importance of the new requests must be weighed against existing bookings scheduled, and where necessary, changes made. This results in further negotiations with production departments, and possible cancellation or postponement of existing bookings. Five days before the machines are due to be used copies of the 'provisional' schedules are sent out to a number of people connected with the use of VT machines. Any subsequent changes must be relayed to these people. Three days later the schedules become 'revised' and more copies are sent out. After this only essential changes are made and the 'final' schedules are distributed the day before the machines are to be used.

The distribution of schedules cannot in reality be this regular, because the recording assistants do not normally work at weekends, and unavoidable occurrences (such as machine breakdowns, or major national events causing transmission reorganisations) make the whole scheduling process somewhat less regular than indicated. The complete scheduling process can be summarised as shown in Fig. 2.

Vol. 5
1977

A SCHEDULING SUB-SYSTEM

A computerised Management Information System will normally comprise of: an information store or data base containing stored data relevant to the management function in an easily accessible form; a means of maintaining this data base by inputing new or altered data; the ability to make certain calculations and processing runs using this data; and a method of speedily obtaining the

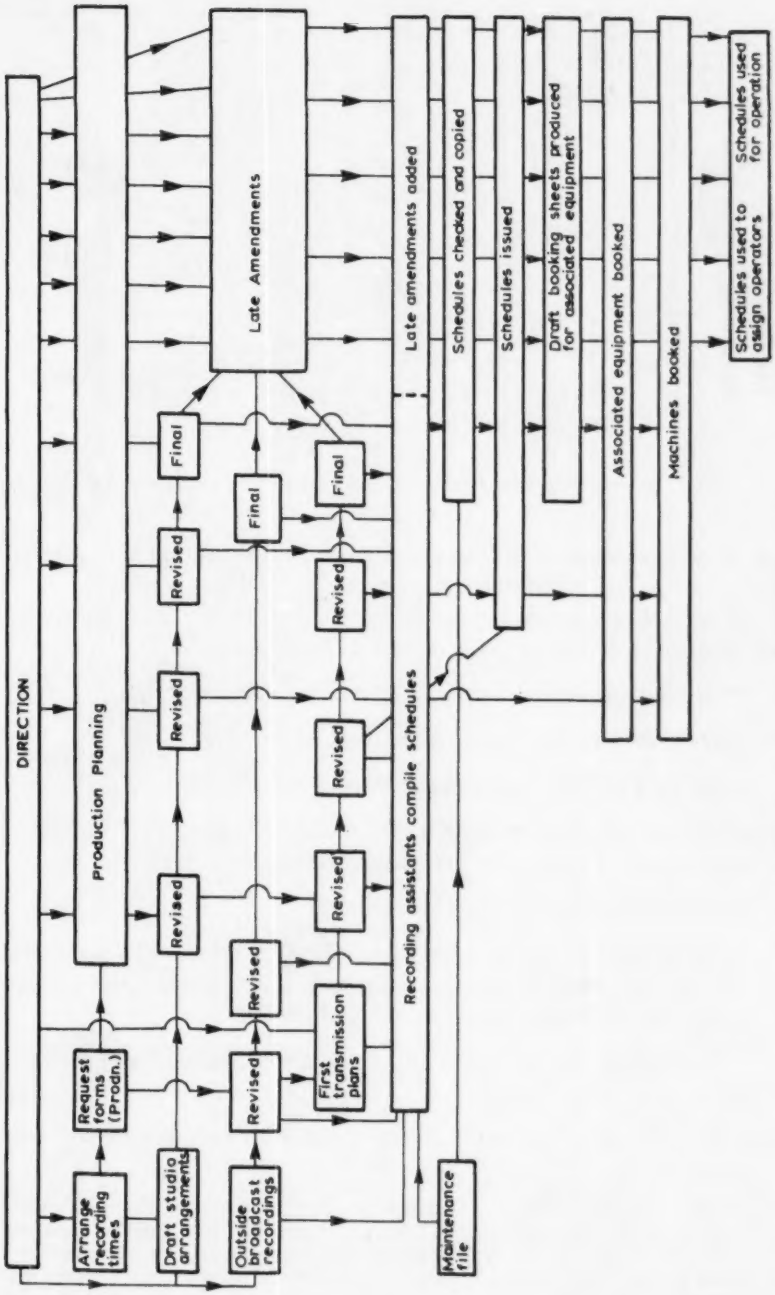


FIG. 2. *A summary of the scheduling process.*

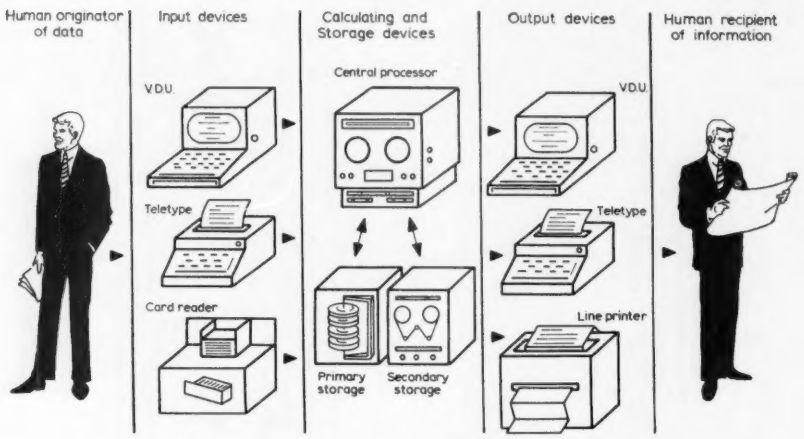


FIG. 3. *Showing the general structure of a Management Information system*

results of such calculations and transmitting them to the relevant management. Figure 3 shows this in diagrammatic format.

(A full account of this research[1] can be found elsewhere [6].) A scheduling subsystem must be able to perform the following functions:

(a) input booking requests for machine time into a computer file;

(b) alter requests as necessary after they have been entered;

(c) delete (and possibly later replace) requests from the file;

(d) search for and print details of a particular specified request, or group of requests which share some specified characteristic;

(e) produce initial schedules for the machines;

(f) allow manual alteration of these schedules in cases where subjective decisions are necessary, or where late requests are received which must be scheduled but do not warrant a total rescheduling;

(g) make available at any time statistics of machine loading, time available etc;

(h) alter the constraints and conditions of the scheduling process as needs arise.

[1] In 1972 BBC Television awarded a bursary to one of the authors (Waters) to enable him to study for the M.Phil. degree at Imperial College under the supervision of another (Jenkins). Close cooperation between Waters, Jenkins and the BBC Television Computer Projects Team (Graham and Howells) over the next 2 years resulted in the development of the system described in the rest of this paper.

In the design of the scheduling sub-system it is necessary to produce an information flow which will both satisfy these requirements and fit into the general structure of Management Information Systems as illustrated in Fig. 3.

An illustration of the sub-system designed for this purpose is shown in the form of an information flow diagram in Fig. 4. Inevitably the constraints imposed on the scheduling will change from time to time, and in order to keep the sub-system working efficiently File 2 must be updated as necessary.

One of the first decisions which has to be made with any such system expansion is the balance between the level of automation and human discretion. A balance has to be drawn, in this case, between doing too much by machine (and therefore not allowing sufficient flexibility in the system), and leaving too much to the schedulers (and thereby not reducing their workload sufficiently to justify the cost of the sub-system).

In the solution illustrated it can be seen that the schedulers have complete control over the system, but by doing all the processing necessary to prepare the initial schedules automatically a large proportion of their routine workload is removed. At the same time leeway is available to alter the schedules in the light of unusual circumstances or requirements.

The schedulers communicate with the computer via Visual Display Units, in an on-line mode, and where possible the outputs are returned through these terminals. However, as there is a large amount of information produced with the schedules this particular output would be best directed to a line printer. In order to reduce delays this should be done at some point where the delivery time is not critical. For example a command to initiate the automatic scheduling

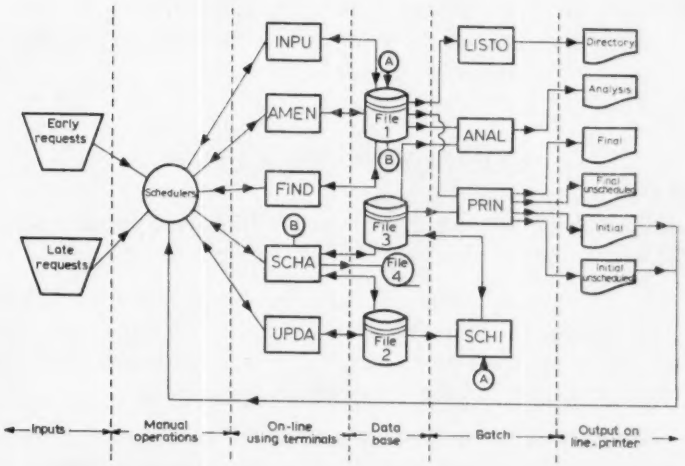


FIG. 4. *The final information flow diagram for the scheduling sub-system.*

run at the end of a working day could make the print-out available before the schedulers arrive at work the following morning.

There are parts of the sub-system which do not involve the use of computers. These include sections where the schedulers receive and transmit documents (either written or verbal), and where they work on these documents by hand. Such operations include the posting and receipt of input data by the schedulers, the manual sorting of these data and the return of the initial and final schedules from the line-printer.

The information involved in those parts of the system not using computers must, to a large extent, be considered as fixed, as no alterations can be made here without affecting the overall information flow in the MIS. However, a distinction may be made between the input information sent to the schedulers by would-be machine users, and the output describing the results of the scheduling runs as returned by the schedulers. In the former case the information handled must be considered as fixed (unless extra work is undertaken to produce additional inputs), while in the latter case the present level of information flow can be considered as the minimum acceptable level.

Additional output information may be made available, but it should be noted that with the use of computers in the sub-system it is very easy to produce large amounts of data. The temptation to distribute the output more widely, and consequently perhaps produce more wastage of time and resources is avoided, with relevant results being distributed only to those whose jobs would be made easier by their receipt.

With the introduction of a computerised scheduling sub-system the actual handling of forms by the schedulers will be reduced considerably. Instead of producing the complete schedules by hand they will only be required to do the initial sorting of data, with any translations necessary to give inputs to the computer, and later examine the printed schedules. All other processing can be done automatically with computer assistance.

Vol. 5
1977

COMPUTER OPERATIONS

As video display units are to be used for all on-line operations there are several factors to be considered:

(a) a VDU works on a 'page at a time' basis (when used as a pure VDU and not as a teletype compatible VDU). Hence a large amount of information should be put on to each screen before it is transmitted. This minimises the delays caused by slow responses, thinking time etc;

(b) overwriting of information on a VDU screen is possible. 'Forms' may therefore be displayed and completed by filling the relevant sections, corrections may be made by overwriting and so on;

(c) there is a limited amount of space on a VDU screen (usually 80 characters by 25 lines, or 56 characters by 18 lines). Coding of information is therefore desirable, both to reduce typing time and also to allow more information to be put on to each screen.

In order to simplify the further examination of this part of the sub-system it can be divided into two parts:

(a) those parts of the sub-system which is concerned with the upkeep and interrogation of data files (Programs INPU, AMEN, FIND and UPDA in Fig. 4);

(b) that part which makes alterations to the automatically produced initial schedules (Program SCHA).

In addition to the schedules themselves, there are several other results which can be best produced on a line printer. These include listings of requests stored in the Master File (to produce a 'directory' of requests), requests left-over after the automatic scheduling run, figures for machine utilisation and tables of spare capacity. Programs LISTO, ANAL and PRIN look after these operations.

THE SCHEDULING ROUTINE

The Program, SCHI, when initiated by the schedulers, reads all the relevant details about requests for machine usage and constraints imposed for the scheduling run from the data files, and automatically prepares the initial schedules. As such it is the most complicated program in the sub-system and can, perhaps, be considered as the most important.

The main characteristics of the scheduling problem to be tackled are:

(a) each job is processed once, usually on one machine;

(b) sometimes a job uses two or three machines working in parallel;

(c) two jobs are sometimes linked so that one must finish before the other can start;

(d) there is a static situation in that all the jobs to be included in the automatic scheduling run are known at the start of the production of initial schedules;

(e) each job has a particular function specified;

(f) there is a basic order of job importance which depends on function, together with additional orders based on other considerations;

(g) there is a variable period of preparation before any job can begin to be processed, which depends among other things on job function;

(h) each machine can handle certain prescribed job functions; but no machine can cope with all functions;

(i) there is an order of machine preferences associated with each job function;

(j) each request can have a different duration;

(k) there may be a fixed or preferred start time associated with each request for machine time;

(l) there may be a fixed or preferred machine associated with each request;

(m) demand for machine time is usually greater than supply;

(n) there are sufficient staff to operate the machines at all times throughout the working day;

(o) the initial schedules will be updated later by using on-line updating runs, in the light of new information available;

(p) the computer capacity available is limited.

Based on these characteristics it is necessary to determine the best means of solution for the automatic scheduling problem. One of the most significant questions which must be answered concerns the choice between optimal and satisficing solutions.

The most obvious of the optimising techniques is complete enumeration, where all possible schedules are examined and the 'best' is determined. However, when the number of possible schedules is examined it can be seen that this approach is not feasible for even a comparatively trivial problem. As an illustration of this the simple case of scheduling 'n' jobs on 'm' machines is examined. If n jobs are to be scheduled on one machine there are $n!$ possible sequences. If 'm' machines may be used the first job of any individual sequence can be put on to any of the alternative machines. In general terms the number of possible schedules for 'n' jobs on 'm' machines is $(n!)^m$. Thus for even a comparatively small problem with four jobs on four machines the number of possible schedules is:

$$(4\ !)^4 \Rightarrow 330{,}000$$

This figure will rise rapidly as the size of the problem is increased. For real problems this sort of enumeration is beyond the scope of even the largest computer to evaluate in a usefully short time. In order to reduce the number of solutions which need to be enumerated, algorithms have been constructed which attempt to 'home-in' on an optimal solution. These algorithms themselves, however, often run into computing difficulties which make them unsuitable for solving real problems.

Gupta [4] summarises the situation by saying, "Scheduling n-jobs on m-machines has been an area of constant research for the past two decades." However, in spite of several research efforts, practical scheduling problems cannot be optimised effectively and efficiently. The reported progress in the field of scheduling is almost insignificant as compared to the practical requirements of the problem". This view is supported by Balas [1] who wrote. "Machine sequencing is one of the most frequently occurring real-world problems, which while theoretically 'solved', cannot be handled for a realistic problem size by any of the existing optimising techniques".

These two comments illustrate the split which exists between 'theoretical' optimising solutions and 'satisficing' solutions to practical problems. At present these two types of solution must be regarded as separate, and there seems very little hope of providing optimal solutions to real problems in the near future. There are, in general terms, only two types of problem for which optimal solutions can be obtained: small-scale problems involving a few machines and jobs, and larger problems in which constraints and restrictions are relaxed to make solution easier.

A real scheduling problem will usually be much more complex than a theoretical formulation. In particular constraints are more rigid, the number of interacting factors is generally greater, the interactions are more complex and subjective decisions are likely to be more relevant. In practice there are two major approaches open for obtaining sub-optimal, satisficing solutions:

(a) Simulation—involving the building of a mathematical model of the system to be investigated. Experiments are then performed on the model and the factors which produce good solutions can be identified. These may then be transferred to the real system. An account of some recent work comparing various loading rules is to be found in a paper by Eilon *et al.* [2].

(b) Heuristics—which are methods of obtaining results by calling on past experience of similar situations. They depend on techniques such as loading rules and priorities.

In this particular scheduling problem there are several factors which must be considered in determining the better approach to a solution. These include:

- the number of jobs to be scheduled;
- the number of machines to be used;
- the number of variables associated with each job;
- the number of variables associated with each machine;

- the complexity of the rules for determining which jobs can be scheduled on which machines;
- the computer resources available.

Because each job is unique, and as there are a large number of possible variations in job characteristics, it was thought that a system of loading rules would be ideally suited to this application. Simulation approaches would prove less satisfactory, due to the complexity of the model which would have to be built.

As the jobs to be scheduled are sufficiently different from each other, a system of loading rules can be devised which will differentiate the requests and put them in some order of priority in which they can be scheduled. This would prove more satisfactory, and less costly in resources, than a more formal simulation approach. It was therefore decided to adopt an heuristic approach, relying on a system of loading rules.

Loading rules for solving the problem

The loading rules devised must depend on those variable characteristics of the requests for machine time which affect the way in which they are scheduled. A lot of the previous work done comparing loading rules for various scheduling situations concentrates on the job duration. For instance Eilon and Pace [3] have shown that in their study of job-shop scheduling the most satisfactory results were obtained by scheduling the jobs in order of increasing duration, i.e. the shortest job is considered first. Further evidence in support of this can be found in another paper by Eilon *et al.* [2].

From the previous descriptions of the particular problem considered in this report it can be seen that the variable characteristics of the requests for machine time include:

(a) job duration;

(b) function—which is equivalent to importance, and also determines the machines which may be used;

(c) whether or not there is a preferred/fixed machine specified for the job;

(d) whether or not there is a preferred/fixed start time for the job;

(e) the arrival time of the requests for machine usage at the schedulers.

These factors all add further complications to the problem considered. As an example, if requests are considered for scheduling in order of function (i.e. most important first) then an early request may be scheduled in a gap which is needed for a later request with 'fixed' characteristics. Conversely if requests are considered in the order of 'most fixed first' (that is considering those requests on

Vol. 5
1977

fixed machines and at fixed times first) all the machine time available may be taken up with comparatively unimportant fixed jobs, while important ones with more flexibility will remain unscheduled.

If either of these straightforward approaches are used the most satisfactory results are unlikely to be obtained. This is true even though the criteria chosen for determining the quality of a schedule (i.e. for deciding what constitutes a 'good' or 'bad' schedule) will inevitably have a considerable effect on the loading rules adopted in any practical situation.

The three most important factors which effected the scheduling in this problem are fixed/preferred conditions, function and duration. A short description is given in Table 1. In this context 'fixed' means that no other time or machine can be used to schedule a request if that specified by the schedulers is not available. 'Preferred' indicates that some other time and/or suitable machine may be used if the specified time and/or machine is not available, but the result will be less satisfactory. 'Any' means that any available time on a suitable machine can be used with equally good results. A 'suitable' machine is one which is capable of handling the request function. This 'fix code' can be used to identify the amount of flexibility of movement of the request. In general terms the higher the fix code, the less flexibility a request has and the more fixed it is. The emphasis placed on the degree to which a request is fixed at a given time or on a given machine will obviously affect the result of the scheduling.

For instance, if requests are scheduled in an order which is not dependent on the fixed conditions, it is possible that early requests which have a greater degree of freedom will be scheduled so that they occupy spaces which are needed for later more fixed requests. It is likely that more satisfactory schedules will be produced if some account is taken of the fix code. In particular it is probably advisable to consider the requests in the order of decreasing fix code for the scheduling run.

Function

As a function also gives an indication of the importance of a request similar considerations apply. Notably if requests are taken for scheduling in an order

TABLE 1. FIXED/PREFERRED CONDITIONS

'Fix code'	Machine	Start time
1	Any	Any
2	Preferred	Any
3	Any	Preferred
4	Preferred	Preferred
5	Fixed	Any
6	Any	Fixed
7	Fixed	Preferred
8	Preferred	Fixed
9	Fixed	Fixed

independent of function, important requests may be left unscheduled due to a lack of free time on the machines.

It is a straightforward matter to encode functions so that an increasing or decreasing function number is associated with increasing or decreasing importance. Then the best results may be expected when the most important requests are scheduled first. If this is not done (and bearing in mind that demand for machine time is greater than supply) all the available machine time may be taken with comparatively unimportant jobs, leaving insufficient space for the more important ones.

Duration

There are basically three possible orders into which the requests can be put with regard to duration:

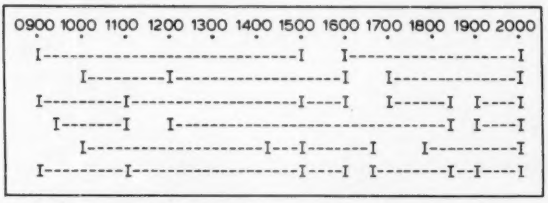
(i) shortest first;

(ii) longest first;

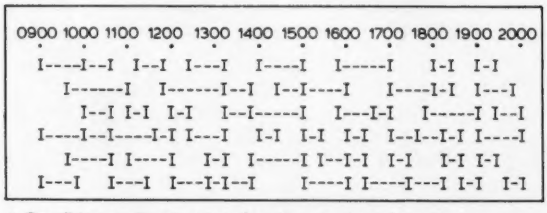
(iii) independent of duration.

In addition there is a further alternative which may be advantageous in cases where there are more requests for machine time than can be scheduled. If the longest jobs are scheduled first, few requests may be scheduled, but efficient schedules may be given. Conversely scheduling the shortest jobs first may produce inefficient schedules, but more jobs could be scheduled. This principle is illustrated in Fig. 5.

Vol. 5
1977



a. Possible results of scheduling the largest jobs first.



b. Possible results of scheduling the shortest jobs first.

FIG. 5. *Illustrating possible effects of scheduling jobs according to duration.*

In an attempt to compromise between these two extremes an alternative was devised which ignored a proportion of the longest requests, and considered the remainder in the order longest first. The proportion of requests ignored rises with the demand for machine time or the number of requests submitted. One drawback with that approach was that the longest requests were always ignored, unless there was an exceptionally low demand for machine time. Even if they were resubmitted for a second scheduling run they were usually ignored.

A set of loading rules was devised and tested to determine which of the three factors considered was most important in the production of a schedule, and which combination of factors produced the schedules which were judged 'best' when measured against certain criteria. From the results obtained in the series of trial scheduling runs it was concluded that the most satisfactory results could generally be obtained by considering the requests for machine time in the order most important, most fixed, longest and earliest arrival first.

This conclusion was based on the assumption that approximately equal importance would be given to each of the criteria by which the schedules could be judged. These are:

(a) always schedule top priority requests;

(b) give a high utilisation of the machines scheduled;

(c) schedule all fixed time requests;

(d) schedule a high percentage of 'preferred machine' requests.

While probable that the approach which proved most satisfactory with the simulated data in the initial tests would satisfy most of these conditions, a second series of tests was undertaken to ensure that this was indeed the case.

CONCLUSION

Not all of the Scheduling Sub-system described in this paper has been implemented. Nevertheless the work done has drawn attention to the power of semi-automatic computer assisted scheduling in the television environment. Indeed the general approach adopted is probably capable of being used in a number of different working environments. The most significant problems encountered by the investigators was the natural reluctance on the part of the existing scheduling staff to conceive of the possibility that a computer-based algorithm could be superior to their own intuitive decision making. Clearly the limited experiments reported here are not conclusive evidence of superiority but it is encouraging to be able to report that some of the earlier suspicion has disappeared.

REFERENCES

1. BALAS E (1970) Machine sequencing: disjunctive graphs and degree-constrained subgraphs. *Naval Res. Logistics Q.* **17**(2), 1–10.
2. EILON S, CHOWDHREY IG & SERGHION (1975) *Experiments with the SI^x Rule in Job-Shop Scheduling in Simulation*, Vol. 24, Part 2, pp. 45–8.
3. EILON S & PACE AJ (1970) Job-shop scheduling with regular batch arrivals. *Proc. Inst. mech. Engrs* **184**(1), 301–310.
4. GUPTA JND (1971) The generalised *n*-job, *M*-machine scheduling problem. *Opsearch* **8**(3), 173–185.
5. WATERS CDJ (1972) The Scheduling of Technical Resources. M.Sc. Report, at Imperial College of Science and Technology, University of London.
6. WATERS CDJ (1976) Computer-Based Facility Scheduling. M.Phil Thesis. Imperial College of Science and Technology, University of London.

ADDRESS FOR CORRESPONDENCE: *Dr JO Jenkins, Department of Management Science, Imperial College, Exhibition Road, London SW7 2BX*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

Optimal Mixed-Model Sequencing for Balanced Assembly Lines

EM DAR-EL
S CUCUY
Technion, Israel Institute of Technology[1]

(*Received July* 1976; *in revised form December* 1976)

This paper describes an algorithm for solving optimally, the mixed-model sequencing problem when assembly line stations are balanced for each model. An optimal sequence is obtained with the minimization of the overall assembly line length for zero station idle time.

The algorithm incorporates two basic steps. The first involves a search procedure that generates all cycle sequences; i.e. sequences having identical 'start' and 'finish' positions and whose work content can be executed within a defined station length. The second step uses integer programming (IP) to determine the number and combination of the various cycle sequences, such that the production demand is satisfied.

1. 5
77

INTRODUCTION

THE mixed-model sequencing problem, first annunciated by Kilbridge & Wester [4], occurs when several models of a particular product are to be assembled on a single assembly line. The authors defined two model launching systems—the *Variable* and *Fixed* Interval Systems, and it is evident that the Fixed Interval Launching System is the more practical of the two [7].

Two basic approaches have appeared in the literature for determining the assembly sequence in mixed-model problems. Tomopolous [5, 6] proposed the use of penalties for various types of inefficiencies that could occur (e.g. idle time, congestion, etc.), so that the subsequent sequence selection is made with the model that incurs the least total penalty over the entire assembly line. The weakness of this method is the arbitrary nature of the penalties used including the physical operating conditions of the assembly line (i.e. permissible station overlaps with adjacent operations, etc.) which have an important influence on the sequence generated. Dar-El & Cother [2] developed a heuristic approach which generates a sequence of models that minimizes the overall length of the assembly line for zero operator idle time. The Dar-El–Cother approach is used

[1] Operations Research, Statistics & Economics, Mimeograph Series No. 178.

in this paper which describes a method that optimally solves the mixed-model sequencing problem for the case where the total assembly time for each model is equally balanced over all stations in the assembly line.

The following notation is used in the paper:

m, total number of different models;

n_j, quantity required of model j, $j = 1,2,\ldots,m$;

N, total number of products required for assembly $= \Sigma n_j$, where Σ denotes the sum from $j = 1$ to $j = m$;

c_j, balanced 'cycle' time of model j in the station considered.

$\bar{c}$, launching rate $= \Sigma c_j n_j / N$;

$l_j, \bar{l}$, length equivalent of c_j and $\bar{c}$ when the models move at a constant speed past the work stations;

L, the station length.

PROBLEM DEFINITION

If the total assembly time for each model is equally balanced over all stations, one need only consider model sequences in the first station since these are repeated identically in each of the other stations in the assembly line. Stations are said to be 'open' or 'closed' [2] depending on whether or not operators are permitted to cross common boundaries. This study applies to closed stations.

The mixed-model sequencing problem can be stated as follows: 'To sequence the N products so that the maximum time the products remain in the station is minimized for zero operator idle time'. The Dar-El–Cother approach, however, requires the transformation of time units into physical length units by assuming that the products are conveyed along the assembly line at a constant speed of 1. Consequently, the time quantities c_j ('cycle' time for model j) and $\bar{c}$ (the launching rate) become length dimensions l_j and $\bar{l}$ respectively and the sequencing problem can be reformulated as: 'To sequence the N products so as to minimize the station length for zero operator idle time'.

Figure 1 illustrates the manner in which the sequenced products flow between the station limits (i.e. boundaries). Models are launched periodically at length equivalent ($\bar{l}$) of $\bar{c}$ and, for simplicity, it assumed that operators move instantaneously from the completion of work on one product to the start of work on the next one on the assembly line.

The station length is L and the starting point is zero. Choosing the first model (No. 3) results in a movement to the right of length l_3. The actual position in the station of the next model (No. 1), which is launched $\bar{l}$ after work on model 3 is started, is found by moving to the left a distance of $\bar{l}$. The sequence and movement of models continue with the only constraint being that station boundaries are not violated. The lower bound (LB) for the station length is the maximum of two estimates [2]:

Vol. 5
1977

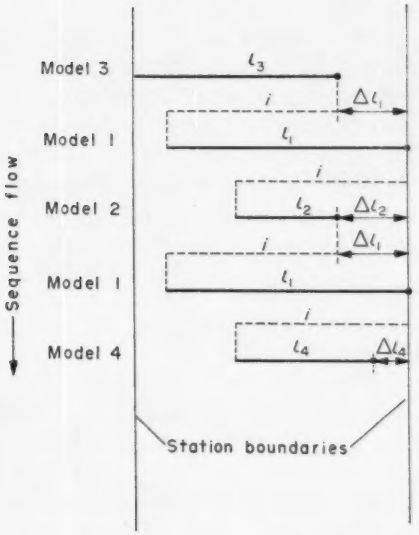


FIG. 1. *Model sequence flow within station boundaries.*

$$LB_1 = \max(c_j)$$
$$LB_2 = 2\bar{l} - \min(c_j)$$
$$\text{thus } LB = \max(LB_1, LB_2)$$

The solution method begins by putting $L = LB$, but it may be necessary to increase the station length to optimally sequence the N products without violating the station boundaries.

Let $\Delta l_j = l_j - \bar{l}$.

Then Δl_j determines the relative position of the completion point of model j by its addition to the completion point of the last model assigned to the sequence. The completion points are shown as heavy dots in Fig. 1, and their relative locations are determined by the appropriate Δl_j values. If $\Delta l > 0$, the movement of the completion point will be to the right; if $\Delta l_j < 0$, the movement of the completion point will be to the left. In order not to violate the lower and upper station boundaries, Δl_j can only move within the distance S, where $S = L - \bar{l}$ (an example is given later). The solution efficiency (E) is defined as:

$$E = (1 - \left(\frac{\bar{l}+S-LB}{LB}\right)) \quad 100\%$$

GENERAL METHODOLOGY

The problem is tackled in two steps—Phase A and B—which are briefly described below.

Phase A—cycle sequences

Starting from any value of P_r(say P_0, $0 \leq P_0 \leq S$), sequences are generated in order to complete cycles. These occur when a model j is found, which, when added as the qth element in the sequence, leads to $P_q = P_0$, i.e. sequences of models are found whose starting and end positions are the same. Such cycle sequences can be repeated at will without violating station boundaries.

The aim of Phase A is to find all possible cycle sequences. At least one cycle sequence must be found in order to proceed to Phase B. If no cycle sequence is found, then the station length (i.e. the value of S) is increased by some increment, ΔS, and the search for cycle sequences reactivated from the beginning.

Phase B—matching the production order

Phase B is the process that determines the number and combination of cycle sequences which satisfies the production order. If no solution is found, the station length is increased by ΔS and the search for cycle sequences is restarted as described in Phase A.

The flow chart in Fig. 2 broadly describes the solution methodology as explained in Phases A and B.

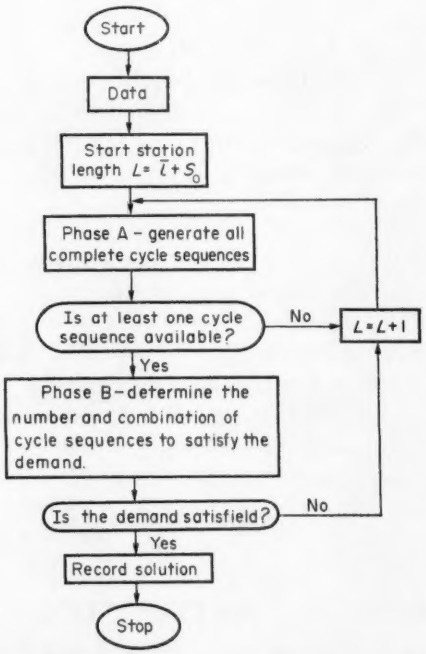


FIG. 2. *Flow chart for solving the mixed-model sequencing problem.*

Convergence. That the algorithm (Phases A and B) will converge can be readily shown. Since the launching rate ($\bar{c}$) was made equal to the grand mean of the cycle times taken over all the production demand (see definition of $\bar{c}$), it follows that $\Sigma\, \Delta l_j.n_j = 0$. Consequently, the entire production requirement can be looked upon as if it were a single 'complete cycle sequence' for which a large, but finite number of solutions exist comprising different station lengths (depending on how the sequence is generated). The optimal solution(s) would then be the sequence(s) associated with the minimal station length.

The search procedure. The systematic search for cycle sequences is important for both control and computation time. Elements (i.e. models) are selected according to any convenient ranking order. The algorithm systematically generates all existing cycle sequences which may occur in two ways:

(i) as complete sequences whose start and end 'P_r' values are zero, and

(ii) as complete sequences whose start and end 'P_r' values are $\neq$ zero. These cycle sequences occur as subsets of those in (*i*) above and constitute an internal loop within the larger sequence. Once identified, the algorithm must take the search process out of the loop to complete a cycle sequence of the Type (*i*) above. These internal loops will themselves be generated later as complete cycle sequences of the type (i) above and are therefore transposed and recorded as such.

A record is maintained of the 'P_r' values of all elements in completed cycle sequences. Thus, in any current sequence generation, if a 'P_r' value obtained is identical to any in the record, then it is clear that a new complete sequence has been found since all P_r values in the record eventually lead to a terminal P_r value of zero. This step considerably reduces the computation time whilst maintaining all the features of an exhaustive search.

Example. Consider a sequencing problem involving 4 models whose Δl_j are as follows:

Model	A	B	C	D
Δl	4	−3	6	−5

For optimal station length

$S = 6$

It is convenient to rank the models in the following manner:

Model	Ranked order
C	$\Delta l_1 = 6$
A	$\Delta l_2 = 4$
D	$\Delta l_3 = -5$
B	$\Delta l_4 = -3$

Figure 3 illustrates the 'Δl movement' diagram and is self-explanatory; it is derived in conjunction with a decision tree search, which is not shown here.

Table 1 summarized the results of the complete cycle sequences obtained in the search process. Some of these can be cancelled since they are, in fact, multiples or combinations of other complete cycle sequences as indicated in the lower part of Table 1.

Thus only cycle sequences 1, 3, 4 and 6 are basic and these are the ones carried over into Phase B. Each of these four sequences may be represented by a vector A_i ($i = 1,2,3,4$) of length m.

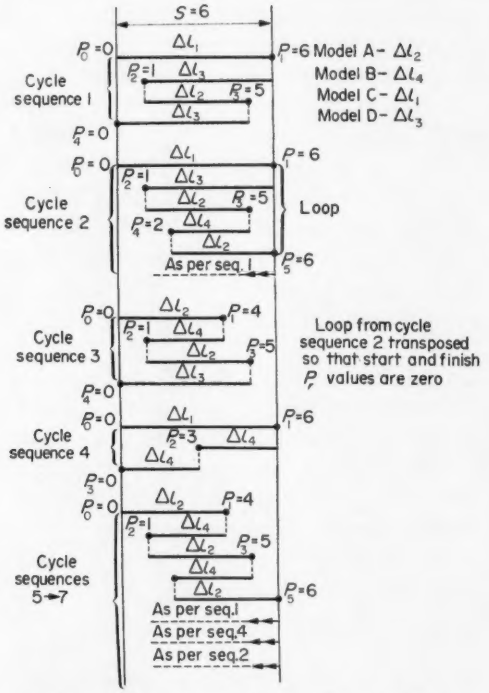


FIG. 3. *Illustrating the 'Δl movement' diagram.*

Phase B—the combination of cycle sequences

The problem of determining the number and combination of cycle sequences to match the demand, can be described as an integer program (IP). If an optimal solution is not found, the station length is enlarged by an increment ΔS and the process of cycle sequence generation recommences from the beginning (i.e. Phase A).

An illustration of phase B. Consider the example where:

$n = 4$ models i.e. $j = 1,2,3,4$.

Vol. 5
1977

TABLE 1.

	Complete cycle sequence No.													
	(1)		(2)		(3)		(4)		(5)		(6)		(7)	
Complete cycles include models	Model	Place	Model	Place	Model	Place	Model	Place	Model	Place	Model	Place	Model	Place
	C	6	C	6	A	4	C	6	A	4	A	4	A	4
	D	1	D	1	B	1	B	3	B	1	B	1	B	1
	A	5	A	5	A	5	B	0	A	5	A	5	A	5
	D	0	B	2	D	0			B	2	B	2	B	2
			A	6					A	6	A	6	A	6
			D	1					D	1	B	3	D	1
			A	5					A	5	B	0	A	5
			D	0					D	0			B	2
													A	6
													D	1
													A	5
													D	0
ΣA	1		3		2		—		4		3		6	
ΣB	—		1		1		2		2		4		3	
ΣC	1		1		—		1		—		—		—	
ΣD	2		3		1		—		2		—		3	
Remarks:	OK		Cancel Set=(1)+(3)		OK		OK		Cancel Set=2 × (3)		OK		Cancel Set=3 × (3)	

$N_j = 50; 40; 60; 50.$
$\Delta l_j = 1{\cdot}6; 0{\cdot}6; -1{\cdot}4; -0{\cdot}4.$
$S_0 = \max(\Delta l_j, j = 1,2,3,4) = 1{\cdot}6$

Applying Phase A yields the following model quantities obtained from 5 basic cycle sequences:

Vectors

Model	$\bar{A}_1$	$\bar{A}_2$	$\bar{A}_3$	$\bar{A}_4$	$\bar{A}_5$
1	1	1	1	0	0
2	2	1	0	3	2
3	2	1	0	1	0
4	0	2	4	1	3

The IP equations are then expressed as follows:

$$\begin{aligned} x_1 + x_2 + x_3 &\leq 50 \\ 2x_1 + x_2 + 3x_4 + 2xz &\leq 40 \\ 2x_1 + x_2 + x_4 &\leq 60 \\ 2x_2 + 4x_3 + x_4 + 3x_5 &\leq 50 \end{aligned}$$

with the objective function:

$$\max(x_1 + x_2 + x_3 + x_4 + x_5)$$

where $x_1 \leq 0, x_1$ — integer, $i = 1,2,. . .,5$

Using a standard IP computer program yields no optimal solution, hence enlarge the station length by $\Delta S = 0{\cdot}1$.

$$\text{i.e. } S_1 = S_0 + \Delta S = 1{\cdot}6 + 0{\cdot}1 = 1{\cdot}7.$$

Repeating Phase A produces the same results as before and the station length is increased once more, i.e. $S_2 = 1{\cdot}8$.

Phase A yields an additional cycle sequence, $\bar{A}_6 = [2 \quad 0 \quad 2 \quad 1]$ and the IP produces the following optimal solution:

$$20\,\bar{A}_1 + 10\,\bar{A}_3 + 10\,\bar{A}_6$$

whose model totals match the respective demands.

Vol. 5
1977

DISCUSSION AND CONCLUSIONS

While the efficiency of the solution naturally depends on the input parameters, the algorithm will nevertheless find an optimal sequence of models for zero station idle time. The algorithm is clearly meant to be used for the design and planning of new mixed-model assembly lines. When changes in demand for production quantities occur in an existing assembly line and present station lengths are imposed as input constraints, the algorithm may not achieve the desired production (but neither can any other known sequencing method guarantee the desired production for the condition of zero idle time. When station lengths are specified as a constraint, it may be desirable to develop tables that give optimal model combinations and sequences for the specific station length. The number of such combinations for each station length can be very large and the production management may find it convenient to select one that best matches its marketing forceast (within some specified tolerance).

The algorithm is inappropriate for solving sequencing problems in which unequal station balances are likely to occur, ie. due to technological constraints For example, when similar tasks on each model must be done on the same station, or else, when tasks requiring a certain type of skill must be performed in one station, then such conditions could result in large imbalances between the work assigned to each station (i.e. for each model) even though the station loads per shift are the same for all stations. When technological constraints do not apply, efficient station balances can be achieved with a high degree of reliability [1] and this algorithm would be very appropriate for producing model sequences.

The computer program developed for Phase A [3] is simple, fast, and does not require a large memory. It is expected to handle problems of up to 10 models, though this has yet to be confirmed statistically. Any standard IP program can be used for Phase B and these are expected to readily handle most industrial problems.

Finally, this paper describes the first successful attempt to optimally solve the mixed-model sequencing problem, even though the requirement for equal station balances for each model somewhat restricts its wider application.

REFERENCES

1. DAR-EL (MANSOOR) EM (1973) MALB—a heuristic technique for balancing large single-model assembly lines. *AIIE Trans.* **5**(4), 343–356.
2. DAR-EL EM and COTHER RF (1975) Assembly line sequencing for model mix. *Int. J. Prod. Res.* **13**(5), 463–477.
3. CUCUY S (1974) *The Mixed Model Assembly Line Balancing Sequencing Problem.* Unpublished M.Sc. Thesis, Technion—Israel Institute of Technology.
4. KILBRIDGE M and WESTER L. (1963) The assembly line model-mix problem. *Proc. 3rd Int. Conf. on Operations Research,* (Ed. DUNOD) English Universities Press, Paris.
5. THOMOPOULOS NT (1968) Line balancing sequencing for mixed-model assembly. *Mgmt. Sci.* **14**, B-59–B-75.

6. THOMOPOULOS NT (1970) Mixed-model line balancing with smoothed station assignments. *Mgmt. Sci.* **16**, 593–603.
7. WILD R (1972) *Mass-Production Management*. John Wiley, London.

ADDRESS FOR CORRESPONDENCE: *Professor EM Dar-El, Technion, Israel Institute of Technology, Faculty of Industrial and Management Engineering, Haifa, Israel.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

Memoranda

Game Theory and Taxation

INTRODUCTION

IN A world of uncertainty, in which the probabilities of occurrence of states of nature are unknown, cardinal utility theory is not always applicable. Instead, game-theoretic strategies may be considered, be it in a less advanced way. The optimal solutions to several types of game may be influenced by the introduction of taxation. In this short paper, we shall focus upon the consequences of different fiscal loss offset regimes on the results of a zero sum game, a game against (or with) nature, and a Nash cooperative game[1]. As the presence of uncertainty is an essential feature in game theory, considerations about offset of loss are a likely research topic.

In practice, there will nearly always exist some asymmetry in the taxation of profits versus losses, by that the government shares less in losses than in profits. Loss compensation will, generally, be less than complete, because

- the carry-back and carry-forward terms may be exceeded;
- costs of interest are inherent in the delay of the realization of the loss offset;
- a differential (lower) tax rate may apply to losses than to profits.

If the fiscal regime taxes positive pay-offs not symmetrically as compared with negative pay-offs, we shall speak of non-neutral taxation. By this, taxation may *a priori* be said to be neutral if full loss offset is allowed for. In the latter case, the *optimal* game solutions which are valid in a taxless world are *not altered* by the introduction of a tax[2]. Our players of the game are supposed to be entrepreneurs, who have command over several strategies. We concern ourselves only with a corporate income tax. We now examine the consequences of introducing a non-neutral tax for optimal game behavior.

In a zero sum game the pay-off to one entrepreneur (duopolist) is, by definition, the negative of the pay-off to the other entrepreneur. Each strives for a maximal pay-off, by which the strategies have maximin-minimax properties [3, chapter 6]. If the game is strictly determined, i.e. a saddle point solution is to be reached, then a non-neutral tax may disturb the original equilibrium in that the game loses its zero sum character: the sum of positive and negative pay-offs, which are asymmetrically taxed, is not zero. The same is true for a non-strictly determined zero sum game, i.e. where no saddle point exists. Here, a solution is to be found by mixing strategies. The originally optimal mixture is also affected by a non-neutral tax.

So, this tax transforms—strictly as well as non-strictly determined—zero sum games into non-zero sum games, by which the original equilibria may be disturbed.

In games against nature, the entrepreneur faces a pay-off structure where the elements depend on different states of nature. As we already stated, the probabilities of occurrence of these states of nature are unknown. In this case, the entrepreneur has to determine his course of action, i.e. to choose an alternative, according to some strategy (criterion) which he has to formulate in advance.

We can study the five well-known strategies In the context of a non-neutral tax, viz. [7, 8]:

[1] We might also study the consequences for optimal game behavior of a differential taxation amongst firms, caused by applying different depreciation rules, deductibilities of financial costs, etc. We confine ourselves in this paper only with differential taxation amongst positive and negative pay-offs, because the analyses and conclusions are quite analogous.

[2] Assuming that the fiscal authority does not participate as a player.

- the Laplace criterion;
- the Wald criterion;
- the maximax criterion;
- the Hurwicz criterion;
- the Savage criterion.

According to the Laplace Insufficient Reason criterion, the 'mean' value of the pay-off elements per alternative must be computed, assuming that the occurrence of all states of nature is equally likely. The alternative with the highest 'mean' value is then selected. By introducing a non-neutral tax, the negative pay-offs tend to become relatively heavier burdens, diminishing the 'mean' values of the alternatives more drastically.

Application of the Wald maximin criterion involves searching for the minimal pay-off per alternative, and the one with the highest minimal pay-off is selected. Choices resulting from application of this criterion are invariant against a non-neutral tax. The minimal pay-off elements, of which the maximal one is to be chosen, may change in relative magnitude, however, not in relative position (assuming monotonicity in the tax relationships).

The maximax criterion implies selection of that alternative which exhibits the highest possible pay-off element. The same conclusion of invariancy holds in this case (the maximal pay-off element remains, after a tax levy, the largest possible amount).

However, when a linear combination of these latter two strategies is applied, then a non-neutral tax may have a disturbing influence on the originally optimal choice. According to this Hurwicz criterion, the average value of the highest and lowest pay-off elements per alternative is computed, using an 'index of optimism' as a weight. If this coefficient equals zero, the maximin criterion is obtained, while a unity value degenerates to the maximax criterion. Again, the alternative with the highest average value is chosen. In the example below, we fix the value of the 'index of optimism' on 0·5. Introducing a non-neutral tax implies a tendency to favor alternatives which, before tax, have a low variability of pay-off values (assuming that non-positive elements are included). Just as in the Laplace case, alternatives containing relatively large loss elements become less favorable.

According to the Savage regret criterion, any decision is compared to the decision which would have been made if the entrepreneur had known which state of nature would have taken place. The difference in pay-off inherent in these two decisions, measures the degree of regret (in absolute values). The maximal degree of regret is computed for all alternatives, of which that alternative with minimal regret is selected. The choice between alternatives according to this criterion may also be sensitive to a non-neutral tax. The latter causes a reallocation of regrets, possibly resulting in a switch to other alternatives, where the regret effect is relatively mitigated. By a zero loss offset, the degree of trauma concerning a negative pay-off element is less affected than that concerning a positive element.

Thus, alternatives containing the former elements become relatively less favorable.

An example

The conclusions stated above, are strikingly of a probabilistic character, because of the casual nature of game theory. This prevents us from drawing general conclusions. It may be useful to illustrate the effect of a non-neutral tax by the example in Table 1(a), where the pay-off matrix has four states of nature (I, II, III and IV) and the pay-offs refer to five alternatives (A, B, C, D and E).

The selected alternatives according to the criteria are as follows:

Laplace criterion, alternative A;
Wald criterion, alternative B;
Maximax criterion, alternative C;
Hurwicz criterion, alternative D;
Savage criterion, alternative E.

Now suppose, that after the introduction of the non-neutral tax (at 50% of positive pay-offs and zero loss offset) the matrix changes into Table 1(b). The optimal choices now become:

Laplace criterion, alternative B;
Wald criterion, alternative B;
Maximax criterion, alternative C;
Hurwicz criterion, alternative B;
Savage criterion, alternative B.

It follows—in this special example—that the alternative with a rather stable, non-negative pay-off structure (alternative B) is by far the most favorite course of action. The other alternatives chosen for the case of zero non-neutral taxes have lost much of their attractiveness, irrespective of the criterion applied, and alternative C remains optimal if the maximax criterion holds.

Vol. 5
1977

TABLE 1.

		(a) No tax					(b) Non-neutral tax of 50%			
		States of nature					States of nature			
		I	II	III	IV		I	II	III	IV
Alternatives	A	2	6	−3	8	A	1	3	−3	4
	B	0	6	4	2	B	0	3	2	1
	C	−8	14	6	−2	C	−8	7	3	−2
	D	−4	12	0	0	D	−4	6	0	0
	E	−4	8	4	2	E	−4	4	2	1

Nash cooperative games

Finally, we consider a Nash cooperative game. This is a non-zero sum game. Again, we assume there are two entrepreneurs, who may coordinate their actions such as to earn some positive pay-off in excess of a certain threat pay-off. The latter equilibrium point has maximin-minimax properties. Under a set of assumptions, specified by Nash, a unique solution to this bargaining problem is reached by maximizing the product of excess pay-offs above the threat level [5].

By a non-neutral tax, the original 'bargaining space' may alter in such a way, that another non-congruent negotiation set emerges (compare the graph below). This may imply a different solution to the cooperative game problem.

An example is shown in Table 2 (a) in which the rows and columns represent alternative courses of action for the second and first duopolist, respectively. The first number in the cells is the pay-off to duopolist 2, while the second number indicates the pay-off to duopolist 1.

For both duopolists, choosing the second alternative promises maximal profits, irrespective of the other's action. The corresponding equilibrium point, characterized by the numbers 6, 6, is called the threat point.

However, by cooperative action the total sum can be enlarged. We exhibit the solution to this problem in Fig. 1. The 'bargaining space' is represented by the quadrangle RSTU, that can be derived from Table 2(a). The originally optimal solution is P, where the Pareto optimal pay-off frontier is touched by the highest possible contour. The contours are rectangular hyperbolas with origin in the threat point (denoted by Q); they represent all combinations of excess profit products having the same value. The solution is to alternate between (21, −6) and (−6, 21). The value of the relevant contour is now 2·25.

The dotted line R^*U^* in the graph represents the pay-off frontier which would occur if taxation is neutral (assuming a tax rate of 0·5). The optimal solution point P^* would maintain its relative position with respect to the new threat point Q'. The value of the

TABLE 2.

	(a) No tax				(b) Non-neutral tax of 50%		
	Duopolist 1				Duopolist 1		
Duopolist 2	7, 6	5, 9	2, 3	Duopolist 2	3·5, 3	2·5, 4·5	1, 1·5
	9, 5	6, 6	21, −6		4·5, 2·5	3, 3	10·5, −6
	3, 2	−6, 21	6, 7		1·5, 1	−6, 10·5	3, 3·5

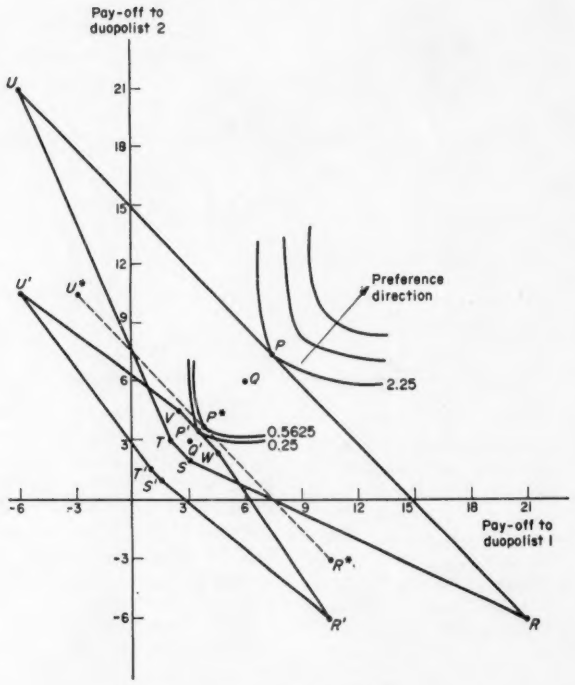


FIG. 1. *Pay-offs for Table* 2.

touching contour would then be 0·5625.

By introducing the same non-neutral tax we assumed before, all positive numbers are halved, while all negative numbers are unchanged, as shown in Table 2(b). Now, the hexagon $R'S'T'U'V'W$ is obtained.

As is clear from the graph, the pay-off frontier $R'W'V'U'$ is now kinked. The optimal solution, denoted by P', is situated on a line segment, connecting (4·5, 2·5) and (2·5, 4·5), implying another course of action than in a taxless or neutral tax world. The duopolists have now to alternate between the above mentioned two points. Clearly, the value of the contour (being equal to 0·25) has been diminished by introducing a non-neutral tax. Note, that the threat point guarantees, after tax, even larger total profits than application of the for tax adjusted original cooperative strategy.

SUMMARY AND CONCLUSIONS

In a world with uncertain pay-off structures, loss offset considerations in relation to optimal game behavior are becoming a relevant research topic. For the types of game treated in the present paper, non-neutral taxation (i.e. asymmetric taxation as to positive versus negative pay-offs) may disturb the originally optimal game solutions, except when the maximin and maximax criteria are applied (in games against nature). As might be expected, alternatives containing relatively large loss elements become less attractive as compared with the taxless world solutions, if the loss offset possibilities are, in effect, less than complete.

REFERENCES

1. HARSANYI JC (1956) Approaches to the bargaining problem before and after the theory of games: a critical discussion of Zeuthen's, Hicks' and Nash's theories. *Econometrica* **24**, 144–157.

Vol. 5
1977

2. HOROWITZ I (1970) *Decision Making and the Theory of the Firm*. Holt, Rinehart & Winston New York.
3. INTRILIGATOR MD (1971) *Mathematical Optimization and Economic Theory*. Prentice-Hall, Englewood Cliffs, NJ.
4. NASH JF (1950) The bargaining problem. *Econometrica* **18**, 155–162.
5. NASH JF (1953) Two-person cooperative games. *Econometrica* **21**, 128–140.
6. VON NEUMANN J and MORGENSTERN O (1944) *Theory of Games and Economic Behavior*. Princeton University Press, Princeton.
7. SCHNEEWEISS H (1967) *Entscheidungskriterien bei Risiko*. Springer, Berlin.
8. TISDELL CA (1968) *The Theory of Price Uncertainty, Production, and Profit*. Princeton University Press, Princeton.

Piet W Moerland
(*November* 1976)

Institute for Fiscal Studies
Erasmus University Rotterdam
Burg. Oudlaan 50
Rotterdam
The Netherlands

ol. 5
977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain.

Chance Constraints and the Beta Distribution

ONE METHOD of solving chance constrained programming problems is to replace each probabilistic constraint by its certainty equivalent. This note discusses the nature of the certainty equivalent when the random variable in the chance constraint follows a beta distribution.

Let us consider the chance constraint

$$P(X \geq R) \geq \alpha \qquad (1)$$

where P represents probability, X is a decision variable, R is a random variable following a beta distribution with parameters a, b, u, and v, and α is a given probability limit. The probabilistic constraint may be expressed in terms of the cumulative distribution function as

$$I_{(X-a)/(b-a)}(u+1, v+1) \geq \alpha, \qquad (2)$$

where

$$I_y(p,q) = [1 + \beta(p,q)] \int_0^y t^{p-1}(1-t)^{q-1}\mathrm{d}t$$

is the incomplete beta function ratio. Since the cumulative distribution function is monotonically increasing, we may obtain the certainty equivalent

$$X \geq a + (b-a)\, I_\alpha^{-1}(u+1, v+1), \qquad (3)$$

where $I_\alpha^{-1}(u+1, v+1)$ is the inverse beta function.

Chance constrained programming results in excess resources that will not be used on the average. The mean of the beta distribution is given by

$$E(X) = a + (b-a)(u+1)/(u+v+2). \qquad (4)$$

Therefore, the proportion of resources in excess of the expected value is $X - E(X)$.

In order to study the effect of changing either the standard deviation of the distribution or the probability limit on excess resources, we may follow the approach of Naslund [1]. Since the standard deviation of the beta distribution is directly proportional to its range, sensitivity analysis may be performed on the range $(b - a)$ instead. When the equality holds in (3), the partial derivative of the RHS with respect to $(b - a)$, which is the inverse beta function $I^{-1}(u + 1, v + 1)$, indicates the proportion by which the excess capacity increases if the range is increased. If the equality does not hold in (3), the partial derivative merely indicates the amount by which the lower limit of the capacity is increased.

The effect of changing the probability limit on excess capacity may be similarly studied. When the equality holds in (3), the partial derivative of the RHS with respect to α yields

$$(b - a)\, \partial I_\alpha^{-1}(u + 1, v + 1)/\partial\alpha,$$

which is the amount by which the proportion of excess capacity increases when the probability limit is increased. But, since there is no closed form expression for the cumulative distribution function of the beta distribution, the expression for the partial derivative of the inverse beta function cannot be calculated. However, its physical interpretation may be easily seen by computing the numerical value of

$$[I^{-1}_{\alpha+\Delta\alpha}(u+1, v+1) - I_\alpha^{-1}(u+1, v+1)]$$

when α increases to $(\alpha + \Delta\alpha)$.

REFERENCE

1. NASLUND B (1967) *Decisions under Risk*. Economic Research Institute, Stockholm School of Economics.

BP Lingaraj*
H Wolfe
(*December* 1976)

**Indian Institute of Management*
33 *Langford Road*
Bangalore 560 027
India

OMEGA, The Int. Jl of Mgmt Sci., Vol, 5 No. 3, 1977. Pergamon Press. Printed in Great Britain

Bias and Variance in Bowman's Managerial Coefficient Theory

BOWMAN'S managerial-coefficient-theory [1] provides a conceptual framework by which to characterize recurrent decisions and use that characterization to improve decision making. Bowman summarized the four points of his theory as follows:

1. In their decision making behavior, managers and/or their organizations can be conceived of as decision-rule coefficient estimators.

2. It is the variance in decision making that hurts (more than bias) due to dish shaped criteria surfaces.

3. A decision rule with mean coefficients estimated from management's behavior should be better than actual performance.

4. It may be better than a rule with coefficients supplied by traditional analysis.

Bowman supported his theory with illustrations from ice cream, chocolate, candy, and paint plants in the Boston area. Numerous studies have been undertaken to establish the theory [2–7].

A verification of Bowman's second point is important because if economic inefficiency is primarily caused by the variance, then using the role must reduce the variance and increase efficiency. This increase in economic efficiency will occur for any reasonable rule. Naturally, an optimal rule would be preferred, but a rule based on human judgement would be very effective. The purpose of this paper is to find the condition under which these propositions hold.

Proposition 2 asserts that the costs associated with variance (erratic behavior) are larger than the costs associated with bias (average non-optimal behavior). Bowman defines those concepts as follows:

> "Departures of the decision making behavior of management from the preferred results, in this sense can be divided or factored into two components, one which in the manner of a grand average departing from some preferred figure, we call bias (which causes a relatively small criteria loss due to the dish shaped bottom of the criteria surface), and one which represents individual occurrences of experiences departing from the grand average, we call variance (which causes larger criteria losses due to the individual occurrences up to the sides of the criteria dish shaped surface). It is the latter and more important component which seems to offer the tempting possibility of elimination through the use of decision rules."

Let X be a random variable representing the series of decisions. Assume that a decision X is associated with the cost Y by the quadratic cost function:

$$\text{Cost} = a + bX + cX^2. \tag{1}$$

If X is normally distributed with mean μ and standard deviation σ then:

$$E(\text{Cost}) = a + b\mu + c(\sigma^2 + \mu^2). \tag{2}$$

The minimum cost point is found by taking the partial derivatives of preceding equation and setting them to zero:

$$\mu_{\min} = -b/(2c), \tag{3}$$

$$\sigma_{\min} = 0. \tag{4}$$

The bias Δ is the difference between μ and $\mu_{\min}$:

$$\Delta = \mu + b/(2c). \tag{5}$$

Substituting the definition of Δ into (2) yields:

$$E(\text{Cost}) = a - b^2/(4c) + c\sigma^2 + c\Delta^2. \tag{6}$$

From (6) it is clear that σ and Δ have an equal impact on expected cost. Bowman's proposition then is an assumption about the characteristics of decision makers rather than a functional relationship. It is true only if the bias in the decision variable is less than one standard deviation ($\Delta < \sigma$). These results generalize to n decision variables.

The production scheduling problem has received much attention (see references) and numerous models have been presented, and perhaps the best known formulation is by

Holt, Modgiliana, Muth & Simon [4] which assumes a quadratic cost environment. Optimal or linear decision rules can be found to minimize the total cost.

Research in production scheduling has shown decision rules effectiveness to be affected by the levels of performance feedback given and the nature of the cost function [3]. To make this experiment a rigorous test of Bowman's proposition, the experimental design incorporates three levels of performance feedback and two different cost functions.

The experimental subjects made production and workforce decisions via timesharing computer. The subjects were business majors from a required undergraduate class in computer-aided forecasting and decision making. Participation in the experiment was a course requirement.

After 16 practice periods and 2 weeks delay, the subjects made decisions for 12 periods. All the cost and inventory calculations were done by the computer. The subjects were randomly assigned to three treatments of performance feedback within both the linear and quadratic cost environments. All subjects were given sales forecasts for the next three time periods. Based on these forecasts, their current inventory position, their current work force size, and worker productivity, the subjects made work force and production volume decisions. All subjects were then fed back the actual sales, current inventory, actual cost, and average cost incurred to that point in the simulation. Subjects in treatment 1 received only this information. Subjects in treatment 2 also received feedback on the units of inventory error, work force change and over/idle time caused by their decisions. Subjects in treatment 3 received all information given to those in treatments 1 and 2 plus a cost breakout for the three factors in treatment 2. The subjects did not know the cost coefficients nor cost structure.

For each condition, regression equations for the work force and production volume decision were found. In the quadratic cost conditions Δ was calculated relative to the optimal rule. In the linear cost conditions Δ was calculated relative to an optimal rule based upon a quadratic approximation of a linear cost function. For all levels of performance feedback in both linear and quadratic cost conditions, Δ was much less than σ, as shown in the table. We conclude that Bowman's second proposition holds. Variance is the major contributor to the inefficiency.

TABLE 1. EXPERIMENTAL RESULTS FOR Δ AND σ

	Information levels		
	(a) Quadratic cost		
Workforce rule	1	2	3
Δ	−20·9	−18·7	−21·4
σ	341	327	344
Production rule			
Δ	−26·6	−18·4	−24
σ	1151	1128	1172
	(b) Linear cost		
Workforce rules			
Δ	−26·7	−50·4	−36·7
σ	446	351	335
Production rule			
Δ	−30·8	−136	−90·9
σ	1217	1183	1208

Other experiments in production scheduling contain cost information which bears on Premise 2. Another experiment of the same design as the foregoing calculated the cost associated with bias and variance directly. The procedure was: (1) Find the characteristic rule and simulate to find the cost of its use. (2) Bias cost = rule cost − optimal cost, (3) Variance cost = actual cost − rule cost.

Bias cost always exceeded variance cost. Moskowitz and Miller [6, 7] reached similar conclusions when using the latter procedure.

In conclusion, a key proposition of Bowman's managerial coefficient theory is that variance is a larger source of economic inefficiency than bias. Assuming a quadratic cost structure, this proposition requires that the standard deviation of the decision variable be greater than the mean bias. The data from our production scheduling experiment show that the proposition holds.

All points of Bowman are now empirically demonstrated. The use of a decision rule may now be advocated because of the economic importance of reducing erratic behavior in addition to arguing for the closeness of the rule to an optimal rule. Hence, the theory now has two strong arguments for its utilization to improve decision making.

Vol. 5
1977

REFERENCES

1. BOWMAN EH (1963) Consistency and optimality in managerial decision making. *Mgmt Sci.* **9**(2), 310–321.
2. CARTER PL and HAMNER CW (1975) A comparison of alternate production management coefficient decision rules. *Decision Sci.* **6**(2), 324–336.
3. EBERT R (1972) Environmental structure and programmed decision effectiveness. *Mgmt Sci.* **19**(4), 435–445.
4. HOLT DC, MODIGLIANA F, MUTH JF and SIMON HA (1960) *Planning Production, Inventories, and Work Force*. Prentice-Hall, Englewood Cliffs, NJ.
5. KUNREUTHER H (1969) Extensions of Bowman's theory of managerial decision making. *Mgmt Sci.* **15**(8), 415–439.
6. MOSKOWITZ H (1974) Regression models of behavior for managerial decision making. *Omega* **2**(5), 677–690.
7. MOSKOWITZ H and MILLER JG (1975) Information and decision systems for production planning. *Mgmt Sci.* **22**(3), 359–370.

William Remus
(*October* 1976)

University of Hawaii
2404 *Maile Way*
Honolulu
Hawaii 96822
USA

ol. 5
977

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 3, 1977. Pergamon Press. Printed in Great Britain

International Register of Theses and Dissertations

The following are obtainable at the Library, Cranfield Institute of Technology, Cranfield, Bedford MK43 0AL, UK:

ANDRIOTIS G (1973) Optimal Relay Station Networks. MSc Thesis, Cranfield Institute of Technology.

BARNARD NR (1973) The Effects of Advertising on the Sales of a Non-Durable Product. MSc Thesis, Cranfield Institute of Technology.

BIRDSEYE BG (1972) Public Transport Operation—A Route Study Technique. MSc Thesis, Cranfield Institute of Technology.

BRENNAN PM (1973) The Application of Computers to the Process Planning of Aircraft Components. MSc Thesis, Cranfield Institute of Technology.

BULLIVANT PE (1973) An Investigation into the Methods of Improving Morale in the Aircraft Industry. MSc Thesis, Cranfield Institute of Technology.

CLARK J (1973) A Study of the Effects of Placing on the Learning and Performance of a Simple Assembly Task. MSc Thesis, Cranfield Institute of Technology.

CLARKSON NS (1971) An Investigation into the Optimisation of Depot Stock Levels for a Manufacturer of Materials Handling Equipment. MSc Thesis, Cranfield Institute of Technology.

COOPER AE (1972) The Assessment of Alternative Strategies for Bus Priority at Signal-Controlled Traffic Intersections. MSc Thesis, Cranfield Institute of Technology.

DEAN C (1973) An Investigation into the Transmission of Diffraction Patterns Through Fibre Optics in Order to Generate Holograms. MSc Thesis, Cranfield Institute of Technology.

DU MEE AMR (1973) The Analysis of Multifunction Control—An Application in Current Motor Vehicles. MSc Thesis, Cranfield Institute of Technology.

GAYTON FM (1972) A Review of Long Range Planning and Forecasting (by Classification and Computer Filing). MSc Thesis, Cranfield Institute of Technology.

HARRISON A (1973) A Material Control System for an Aircraft Company. MSc Thesis, Cranfield Institute of Technology.

HAYWARD RG (1971) John Hoskyns and Company Limited—A Strategy. MBA Thesis, Cranfield Institute of Technology.

HERNANDEZ C (1973) An Analysis into the Inspection Scheme of an Assembly Firm. MSc Thesis, Cranfield Institute of Technology.

HILDERBRANDT MN (1971) The Implications of Possible Improvements in the Utilisation of Coiled Steel in Chrysler United Kingdom Limited (Joint thesis with NJ Quick). MBA Thesis, Cranfield Institute of Technology.

HUSSAIN AI (1972) Simulations of Seaports Applications in Planning and Operations Management. MBA Thesis, Cranfield Institute of Technology.

LATHAM RHR (1971) H Nount and Sons Ltd—An Analysis of the Effect of British Entry into the European Economic Community on the Company, and Recommendations for Future Strategy. MBA Thesis, Cranfield Institute of Technology.

LEGGATE RL (1971) An Investigation of the Management Control System at Hindley Green. MBA Thesis, Cranfield Institute of Technology.

LING JCRE (1971) Daventry—An Overspill Agreement. An Examination of the Movement of People and Industry to Daventry. MBA Thesis, Cranfield Institute of Technology.

LINTON MJ (1971) A Tentative System for the Categorization of Trainer Intervention in T-Groups, and the Study of the Application of this System to the Analysis of Fifteen T-Groups. MBA Thesis, Cranfield Institute of Technology.

LLOYD RG (1972) An Investigation into the Prospects for Expanding the Operations of a British Company into Kenya. MBA Thesis, Cranfield Institute of Technology.

LOVELL WS (1971) A Study for a European Leasing Operation for Mw Brandt's Sons & Company Ltd. MBA Thesis, Cranfield Institute of Technology.

MAGUIRE J (1970) A Computer System for Controlling Raw Material Stock Levels, Applied to the Aircraft Industry. MSc Thesis, Cranfield Institute of Technology.

MURROW KA (1971) Group Technology and its Application to a Tool Room. MSc Thesis, Cranfield Institute of Technology.

NANDA A (1972) A Distribution Study for Associated Deliveries Ltd, Reading (Joint Thesis with DIGENIS G). MSc Thesis, Cranfield Institute of Technology.

OBODO ET (1972) An Investigation for a Comparative Dynamic Performance Test Procedure for Vertical Milling Machines. MSc Thesis, Cranfield Institute of Technology.

POPE KJ (1971) A Study of UK and EEC Agricultural Policies in Relation to the UK Pigmeat Industry, with Emphasis on Grain and Feed. MBA Thesis, Cranfield Institute of Technology.

RANGANATHAN V (1970) Comparison of the Effects of Conventional and Electrolytic Grinding on the Fatigue Strength of Titanium. MSc Thesis, Cranfield Institute of Technology.

SHAMSUDDIN SAO (1971) Stock Control in a Small Manufacturing Company. MSc Thesis, Cranfield Institute of Technology.

SMITH J (1972) The Organisation of Health Care in England with Particular Reference to the Elderly. MBA Thesis, Cranfield Institute of Technology.

SMITH VG (1972) A Corporate Investigation of a Small Firm Presently Marketing Pewterware. MBA Thesis, Cranfield Institute of Technology.

SRINIVASAN SS (1970) Development of a Computer Aided Shop Loading System. MSc Thesis, Cranfield Institute of Technology.

SRINIVASAN V (1970) The Design of a Direct Numerical Control System and its Application to Co-ordinate Positioning. MSc Thesis, Cranfield Institute of Technology.

STROUD PG (1972) The Effect of Control Layout on Man/Machine Performance. MSc Thesis, Cranfield Institute of Technology.

SWEET DC (1973) A Critical Analysis of Existing Methods Used in Control of Stocks by a Medium Sized Company. MBA Thesis, Cranfield Institute of Technology.

TAYLOR PM (1972) An Investigation into Holographic Fringes for Use in Metrology. MSc Thesis, Cranfield Institute of Technology.

THOMPSON JC (1970) An Assessment of the Advantages and Disadvantages of Using either a Full Absorption Cost System or a Marginal Cost System. MBA Thesis, Cranfield Institute of Technology.

VINCENT CJ (1972) The Supply of Dolomite/Magnesite Blocks in the UK—A Report on the Picture from Consett Brickworks. MBA Thesis, Cranfield Institute of Technology.

WALKER RJ (1970) The Optimization of Research and Development Effort. MBA Thesis, Cranfield Institute of Technology.

WEBB CC (1972) The Textile Laminates Market for Polyurethane Foam—A New Product Development. MBA Thesis, Cranfield Institute of Technology.

WILKINSON DJ (1970) A Report on the Performance and Potential of the Export Speciality Food Group Section of Brown & Polson, Export Division. MBA Thesis, Cranfield Institute of Technology.

WILLIAMS JW (1970) The Examination of Specified Companies in the Transport Industry with a View to Recommending to the Board . . . MBA Thesis, Cranfield Institute of Technology.

WOOD RLR (1970) Management vs Tenancy—An Investigation on Behalf of Messrs. Truman, Hanbury, Buxton & Co Ltd. MBA Thesis, Cranfield Institute of Technology.

The following are available at the MIT Microreproduction Laboratory, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02139, USA:

BAYAR I (1973) An Analysis of Communication Mix Models for the Ethical Drug Industry. SM Thesis, MIT Cambridge, MA.

BELL HW, JR (1973) A Portfolio Model for CEDCs—The Incorporation of Social Objectives Into Project Portfolio. SM Thesis, MIT Cambridge, MA.

BERTO FA (1973) An Empirical Study of Security Valuation Models. SM Thesis, MIT Cambridge, MA.

BODDE DL (1973) The Surface Mining of Coal. SM Thesis, MIT Cambridge, MA.

BUSBY RC (1973) A Model for Portfolio Performance Measurement. SM Thesis, MIT Cambridge, MA.

CANN RN (1972) Mathematical Analysis and Optimization of a Two-Level Multiechelon Inventory Distribution System. SM Thesis, MIT Cambridge, MA.

CASSIDY KJ (1973) Post-Merger Management of Acquisitions. SM Thesis, MIT Cambridge, MA.

CHAPMAN DL (with ROEBER RJC) (1970) The Role of Attitudes in the Supply of Graduates to Industry. SM Thesis, MIT Cambridge, MA.

CHASE DA (1972) The Integration of Professionals into Research and Development Organizations. SM Thesis, MIT Cambridge, MA.

CHEVALIER MJ (1970) Choice of Strategy by a French Firm Penetrating the North American Market. SM Thesis, MIT, Cambridge, MA.

COHEN PD (with HEYE J) (1971) Implementation of Management Science Models—An Introductory Training Program. SM Thesis, MIT, Cambridge, MA.

DOBKIN KK (1973) The Team Profile—A Diagnostic Tool for Determining Team Development Needs. SM Thesis, MIT, Cambridge, MA.

FERGUSON TD (1972) Corporate Strategy Consideration of the Phase-Out Joint Venture as a Form of Direct Foreign Private Investment. SM Thesis, MIT, Cambridge, MA.

FONTANET XJ (1973) Planning and Control in a Public Employment Service—Case of France. SM Thesis, MIT, Cambridge, MA.

GALE JD (1971) A Behavioral Science Analysis for Management Information System Design and Implementation. SM Thesis, MIT, Cambridge, MA.

GANESH SR (1971) Values and Interventions—Perceptions of Organizational Consultants. SM Thesis, MIT, Cambridge, MA.

GAPUTIS TC (1973) A Comparison of the Predictive Ability of Some Risk Measures for Determining Future Stock Prices Based on Conditional Expectations. SM Thesis, MIT, Cambridge, MA.

GARMAISE G (1973) A New Proposal for Condiminium Housing in Toronto. SM Thesis, MIT, Cambridge, MA.

GAUGHAN FM (1971) An Econometric Analysis of Marketing Data. SM Thesis, MIT, Cambridge, MA.

HADDADI M (with WEILL G) (1972) Distribution of Steel Products in Algeria. SM Thesis, MIT, Cambridge, MA.

HIATT TA (1972) Trade Between the United States and the People's Republic of China—Implications of the Canadian Experience. SM Thesis, MIT, Cambridge, MA.

HAWKINS WM III (1971) Optimality and Stability Results for Generalized Quadratic Performance Indices and Discrete Linear Control Systems. SM Thesis, MIT, Cambridge, MA.

HEMMELSTEIN JM (1971) The Effect of Residence on the Socialization of the MIT Undergraduate. SM Thesis, MIT, Cambridge, MA.

HOLTHAUS RA (1971) A Time-Sharing Based Information and Accounting System for International Banking. SM Thesis, MIT, Cambridge, MA.

HOTCHKISS JR (1971) A Simulation Study of an Exact Routine for Lot-Size Determination in a Multi-Stage Network. SM Thesis, MIT, Cambridge, MA.

HUNTLEY P (1971) The Effect of Inflation Uncertainty on the Corporate Investment Decision. SM Thesis, MIT, Cambridge, MA.

IYER S (1972) Kogai—Japan's Gross Nuisance Problem. SM Thesis, MIT, Cambridge, MA.

JACOBSON J (1972) Cost of Capital Estimates for the Automotive Industry. SM Thesis, MIT, Cambridge, MA.

JACQUES FM (1971) Motivating Corporate Policy Implementation in Managers of Domestic and Foreign Operations. SM Thesis, MIT, Cambridge, MA.

KANAL P (1971) A Mathematical Model for a Cash Collection System. SM Thesis, MIT, Cambridge, MA.

LANDAU T (1973) The Impact of Conversational Man-Made Machine Decision System on the Portfolio Selection Process. SM Thesis, MIT, Cambridge, MA.

LAW JM (1970) Optimal Capacity Expansion Planning for Production and Distribution Systems. SM Thesis, MIT, Cambridge, MA.

LEVASSEUR RE (1973) A Statistical Analysis for a Marketing Planning Model. SM Thesis, MIT, Cambridge, MA.

LIVINGSTON SJ (1973) Short-Term Capital Market Theory Implications for Mutual Funds. SM Thesis, MIT, Cambridge, MA.

MAGILL MJ (1970) The Influence of Foreign Educational Experience Upon the Transfer of Technology into Ireland. SM Thesis, MIT, Cambridge, MA.

MCLEOD WN (1970) The Impact and Operation of Boston Private Employment Agencies. SM Thesis, MIT, Cambridge, MA.

MENNING WR (1970) The Measurement of Educational Output—An Investigation of the Value Added Concept. SM Thesis, MIT, Cambridge, MA.

MEYER JM (1970) Mixed Ventures Between International Corporations and Government Institutions—The Venezuelan Case. SM Thesis, MIT, Cambridge, MA.

MORRISON KN (1970) A Model of Consumer Response to Innovation Information. SM Thesis, MIT, Cambridge, MA.

MUNOZ MY (1970) An Empirical Study of the Effects of Architectural Change on Organizational Communication Patterns. SM Thesis, MIT, Cambridge, MA.

NAIGLES J (1970) The Partnership Method of Taxing Corporate Earnings. SM Thesis, MIT, Cambridge, MA.

NEALON DC (1970) An Approach to the Establishment of a Certificateless Securities Market. SM Thesis, MIT, Cambridge, MA.

O'CONNOR RW (1971) Use of a Simulation Model as A Decision-Support Tool in the Management of Metropolitan Family Planning Programs. SM Thesis, MIT, Cambridge, MA.

PATTERSON WR (1971) International Ocean Regimes—The Influence of Their Structure and Policy on Corporate Risk. SM Thesis, MIT, Cambridge, MA.

RAND JE (1971) Evaluation of Electronic Cash Registers for Point-of-Sale Data Capture. SM Thesis, MIT, Cambridge, MA.

REGAN MD (1971) Management of Exploration in the Metals Mining Industry. SM Thesis, MIT, Cambridge. MA.

RYBA MW (1971) The Adjustment of Security Prices to a New Listing Information. SM Thesis, MIT, Cambridge, MA.

SCHUMACHER PM (1971) Decision Making Processes of Investment Advisers. SM Thesis, MIT, Cambridge, MA.

TAN KH (1971) Investment in Taiwan—A Look at the Petrochemical Industry and US Involvement in its Growth and Development. SM Thesis, MIT, Cambridge, MA.

VAN LUVEN M (1972) An Intrepreneurial Enterprise Based on an Agricultural Waste. SM Thesis, MIT, Cambridge, MA.

YOSHIMURA S (1972) Corporate Planning in Japan. SM Thesis, MIT, Cambridge, MA.

ZAGER J (1972) EDP Organization in the Municipality. SM Thesis, MIT, Cambridge, MA.

ZAIC GF (1972) The Success and Failure of New Ventures within a Small-to-Medium-Sized Company. SM Thesis, MIT, Cambridge, MA.

The following are obtainable at the Yvette Jacobson Library, Management Centre, University of Bradford, Emm Lane, Bradford, Yorkshire BD9 4JL, UK:

ABRAHAMS IC (1975) An Appraisal of Fundamental Models of Equity Values. MSc Thesis, University of Bradford.

AGUIAR HM (1975) Historical Developments and Applications of Utility Theory. MSc Thesis, University of Bradford.

BARNES PA (1975) The Effect of a Merger on the Share Price of the Attacker. MSc Thesis, University of Bradford.

BLACKMAN CA (1974) Department Store Image—A Case Study. MSc Thesis, University of Bradford.

BUTTERY EA (1975) Corporate Planning in the Multinational Enterprise. MSc Thesis, University of Bradford.

CHAPMAN C (1974) The Principles and Practices of Accounting and Disclosure of Property Companies. MSc Thesis, University of Bradford.

COPE DW (1974) Small Business Finance—A Historic Review and Current Assessment. MSc Thesis, University of Bradford.

DAVIDSON NF (1975) US Banks in the United Kingdom. MSc Thesis, University of Bradford.

DICKINSON DC (1974) Worker Participation—A Review. MSc Thesis, University of Bradford.

EWINS A (1974) Innovative Behaviour and Communications in the Adoption of a Colour Television Set. MSc Thesis, University of Bradford.

GIDENGIL BZ (1975) Social Responsibility and Marketing. PhD Thesis, University of Bradford.

GUY P (1975) A Test of the Financial Life Cycle Theory of the Firm. MSc Thesis, University of Bradford.

HARRINGTON HM (1975) Some Determinants of Aggregate Planning. MSc Thesis, University of Bradford.

HIGGINS JC (1974) A Collection of Publications in Operational Research and Management Science, Management Information Systems and Corporate Planning. PhD Thesis, University of Bradford.

HOLLEY DG (1975) Marketing in the European Economic Community. MSc Thesis, University of Bradford.

LEE L (1974) A Comparative Study of Career Patterns in Physical Distribution in the UK and USA. MSc Thesis, University of Bradford.

LEVINGTON VI (1974) The Marketing and Public Relations Interface in Large UK Companies. MSc Thesis, University of Bradford.

MOLANDER C (1974) A Formal Theory and an Empirical Study of Role Dissonance. PhD Thesis, University of Bradford.

NETHERTON DR (1975) The Effects of the EEC on Marketing Practices Within the Common Market. MSc Thesis, University of Bradford.

PARKINSON JM (1795) The Management of Trade Creditors. MSc Thesis, University of Bradford.

SANDERSON SMcL (1974) The Effect of Multinational Enterprise Upon Market Structure, with Particular Reference to the UK Chemical Industry. MSc Thesis, University of Bradford.

SEDDON VJ (1975) Simulation Study of Management on a Small Dairy Farm. MSc Thesis, University of Bradford.

TEMPLETON R (1975) Profits on Buying into Takeover Bids on the London Stock Market During 1967–1970. MSc Thesis, University of Bradford.

UPEX JR (1975) The Location of Financial Outlets. MSc Thesis, University of Bradford.

WASS B (1975) Design of Postgraduate Production Management Courses. MSc Thesis, University of Bradford.

Information for inclusion in this Section, presented in the appropriate form, should be sent to the Editor of Omega.

Vol. 5
1977

OMEGA

The International Journal of Management Science
Volume 5, Number 4, 1977

Contents

SAMUEL EILON	361	Editorial—A School for Public Policy
PATRICK RIVETT	367	Multidimension Scaling for Multi-objective Policies
RJ BETTS	381	A Review of Inflation Accounting and its Economic Implications
RICHARD M BURTON, BØRGE OBEL	395	The Multilevel Approach to Organizational Issues of the Firm—A Critical Review
R ROTHWELL, J TOWNSEND, M TEUBAL, P SPILLER	415	Some Methodological Aspects of Innovation Research
PAT JOYNT	425	Contingency Analysis for Effective Administration
SAMUEL EILON	437	Management Perspectives in Physical Distribution
NOEL D URI	463	Forecasting: A Hybrid Approach
A DALE FLOWERS, STEPHEN E PRESTON	473	Work Force Scheduling with the Search Decision Rule
RAKESH K SARIN	481	Screening of Multiattribute Alternatives
	491	International Register of Theses and Dissertations
	497	Announcement—A One-Day Conference on Theory and Practice in Management Science

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

EDITORIAL

A School for Public Policy

JUDGING from economic and social trends over the last century it is difficult to escape the conclusion that society is becoming increasingly complex, that the possible effects of business enterprises on the well-being of people and of nations are often too serious to be ignored by government agencies, and that pressures for regulatory measures to control the operations of industry and commerce are constantly mounting. The consequences of such phenomena are manifest in rising public expenditure and inevitably in the number of people employed by the public sector, engaged either in those operating activities which are thought to properly belong to the public domain or in exercising the governmental monitoring and control functions that are regarded as necessary and desirable.

As monitoring and control increase, so does the need to delineate public policy, to determine allocation of resources, to scan alternative courses of action, to assess the impact of short term actions on the prospects of long term well-being. Since vast sums of money are often at stake, involving long lead times in implementation (and a long wait for results of given actions to emerge), it is natural to expect that decision makers in the field of public policy should be properly equipped to deal with problems of great intricacy and be fully conversant with whatever sophisticated analysis is called upon to assist them in their work.

It was therefore with great interest that I read the recent account by Wildavsky, the first Dean of the Graduate School of Public Policy at the University of California in Berkeley, of why and how the School was set up in 1968 and of what has been achieved since [2]. Now, schools of business and public administration had long been established in the US and one can only infer—at the risk of provoking a long and tiresome semantic discussion—that the founders of the School of Public Policy were either dissatisfied with what existing schools could offer, or had something quite different in mind, although what precisely it was that they were dissatisfied with or aiming at is certainly not spelt out. Wildavsky goes even as far as to studiously avoid defining what the term 'public policy' is supposed to encompass, except that "policy analysis is an art form in which problems must tentatively be solved before they can be understood", arguing that "analysis requires creative juxtaposition of resources and objectives until

problems are found that decision-makers can solve" and "only through a model that specifies a desired relationship between manipulable means and obtainable objectives can an analyst make the essential distinction—between a puzzle about which nothing can be done, and a problem for which there may be a programmatic solution".

Such a preamble is bound to raise many semantic problems. I do not know what Wildavsky means by 'a programmatic solution'—if he means programmable, in the sense that an unambiguous procedure can then be set up to handle problems of a given class, then I fail to see the necessity (as opposed to the convenience and resultant efficiency) of such an attribute to the problem solving process, particularly if it is supposed to be a creative one. Also, Wildavsky's definition of a puzzle appears to be associated with some amorphous tangle, whereas I see a puzzle as a clearly defined problem for which the objectives and terms of reference are plainly prescribed for the analyst, whose role is confined to finding a solution to the puzzle and not to question its validity; in other types of managerial problems his possible contribution is obviously of a very different kind, though perhaps not in the sense in which Wildavsky makes his proposition.[1]

Vol. 5
1977

And so, in the absence of a definition of what public policy is, and in view of semantic difficulties that arise in the way the topic is introduced, one needs to turn to other clues in the article. Some are not much help, for example the suggestion that the School should be "constantly repeating this Litany: *What do we teach? Policy analysis. What can students do with it? Many different things*". I know one physics professor who says precisely the some thing, except that he substitutes 'physics' for 'policy analysis', and he sincerely holds that an education in physics is the best preparation for bright young men who want to pursue many diverse careers, including accountancy and the civil service. But the following extracts (with the proviso that the selection is meant neither to be comprehensive nor an attempt to infer what Waldavsky considers important) may give a good indication of his recommended approach and possible methodology:

> "Problem-solving, hopefully—rather than fact-grubbing"—should "predominate".
>
> "Technical competence, statistical supremacy, formal elegance, while all desirable, are not as important as the ability to conceptualize a problem, find an appropriate model for it, develop data appropriate to the task, and emerge with useful information or advice".
>
> "What are the various criteria of choice? How are they to be applied? How can problems of implementation be made part of policy design?"
>
> "Develop people who think of themselves not as problem avoiders, but as problem solvers".
>
> "Emphasize analysis, not subject matter . . . analysis is a stance, not a technique".
>
> "Merely learning facts . . . would defeat the central thrust of a program dealing with problem recognition and solution".

[1] A further elaboration of these points is given elsewhere ([1], p. 139).

In some sense this catalogue of desiderata could be said to be relevant to educational programmes in many disciplines, for example in the education of engineers, designers, architects, town planners. In fact, it would apply to any discipline or occupation in which decision making, problem solving and resource allocation are predominant. There is nothing, therefore, in this list of prescriptions that is exclusively related to public policy.

It is for this reason that I was looking for a rationale for the new school being set up as a separate entity and not as a part of the normal activities that are found in schools of business or public administration. A possible explanation is that even though the approach and the analytical tools, in as much as they are applied, are the same, the problems are not, and some of the differences are evident from the following:

- public administration is concerned with the setting up and manning of procedures that fulfil given aims and obligations, in aligning priorities and in controlling or changing future trends;
- public administration is largely micro, while public policy is aimed at the macro level and often involves a different time scale;
- public policy can rarely be divorced from political considerations and aspirations and hence its practitioners can hardly hope to isolate themselves from the political arena; although political factors obviously affect public administration and the business sector as well, they do so to a much lesser degree;
- businessmen make money, public administrators spend it, public policy planners procure it;
- the more global the approach, the more problems and objectives become diffuse and open ended.

Now, whether such differences justify a separation between schools of public policy and public administration is open to question. If the assertion is made that public policy centres on aims while public administration is only concerned with means, then no doubt many deans of public administration schools will take great exception to such a distinction with respect to the career prospects and aspirations of their students, and even with respect to the programme curricula that should be adopted by their schools. By analogy—so they could argue—such an assertion could lead to the creation of a school for business policy as a separate entity from a school of business administration, a proposition that would be regarded with great suspicion by most academics and practitioners.

Similarly, a distinction between problems concerning public policy and those concerning business is not always easy to sustain. For example, is the building of Concorde a business problem or an exercise in public policy? The construction of a large power station, the closure of a locally dominating plant, the launching of a massive R & D programme—in short all decisions that involve large scale financial and/or manpower resources—are they business or are they societal

problems? The answer must be that they cannot exclusively be regarded as the one or the other; admittedly, the problems are bound to be viewed from different angles, depending on the allegiance of the analyst and the broad objectives that he is commissioned to pursue, but must this fact necessitate different educational programmes? Is the lack of commitment to one particular point of view—be that of the entrepreneur, the consumer, the taxpayer, the public—not an advantage, in that it forces the students, and hence the future analysts, to appreciate the many ramifications that a given set of circumstances and given proposed solutions need to take account of?

One could go further by pointing out the desirability of mutual understanding and argue that for a healthy mixture of views and experiences society should encourage mobility between the public and the business sectors. Too rigid a segregation between the two camps not only deprives each of the enriching experience of the other, but it perpetuates a state of disdain and mistrust, fuelled by an inevitable clash of interests. It is therefore important to allow business executives to move without hindrance into the public domain, and vice versa, not with any naive notion that conflict can thereby be eliminated, but with the hope that it can be better understood and handled. Now, is it more likely that such mobility would be enhanced by the existence of separate schools? Are graduates from a school of public policy likely to be less motivated to go into business and less attractive for businessmen to recruit?

Vol. 5
1977

These are not meant as rhetorical questions; I have not seen any empirical evidence that purports to give the answers, and although I suspect what they might be, I would have liked to have seen an authoritative discussion of the issues involved. It is often said that our attitudes, our arguments, our handling of issues, even our values, are greatly influenced by our education. It is therefore relevant to ask not only what problems a particular educational policy is designed to solve, but also what problems it is likely to create, and it is in this context that a debate on whether education for the public and for the business sectors should be separate would be valuable.

One argument, sometimes advanced, is that the difference between the two sectors is inherent in their separate orientations and hence in the analytical tools that should be employed: the business world is interested in profit, whereas the public sector has many objectives to contend with, so that the optimising approach is the natural methodology for business enterprises to adopt, while satisficing comes naturally to public administrators and policy makers. Such contentions are, of course, totally erroneous and it has been pointed out often enough that most business problems involve multi-goal systems and their handling is very much governed by the satisficing approach.[2] Thus, the differences in methodologies tend to be exaggerated: analytical tools need to be chosen in relation to the problems that emerge, and it is not uncommon to find a particular technique being applied in a variety of circumstances, both in the

[2] See, for example [1, p. 112].

public and in the business sectors. It is not the use of techniques that is the major distinguishing characteristic, but the general milieu that determines the workings of an enterprise or an institution, and while some of the generalisations made earlier are undoubtedly pertinent (for example, attitudes and modes of operations are bound to be affected by whether an organisational entity is expected to make money or to spend it), it is important to realise that there are many major differences to be found between enterprises within the business sector as those postulated between the 'typical' business and the 'typical' public institution.

I am, of course, not opposed to the idea of a school for public policy; I was merely searching for compelling reasons for its justification. But it is rare to find in the educational field compelling reasons for the adoption of a single and universal solution. And it is precisely because dogmatic solutions are both suspect and dangerous that we need variety—in approaches, in design of curricula, and in organisational structures of schools. I am therefore greatly relieved and heartily welcome Wildavsky's cautionary note, not only that each university milieu is different, but also that there should be no "fixation on a single (no doubt idiosyncratic) model" for a school of this type. We learn from experimentation.

SAMUEL EILON
Chief Editor

REFERENCES

1. EILON S (1977) *Aspects of Management*. Pergamon Press, Oxford.
2. WILDAVSKY A (1977) Principles for a graduate school of public policy. *Urban Analysis* **4**, 3–28.

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

Multidimension Scaling for Multiobjective Policies

PATRICK RIVETT

University of Sussex

(*Received November* 1976)

This paper develops and expands the use of multidimension scaling techniques (MDSCAL) as applied in the two separate fields of psychological testing and archaeology to the problem of multiple criteria decision making. Other work by the author published elsewhere shows that it is feasible to use MDSCAL for drawing maps of separate policies using very weak input information from which deductions as to most preferred and least preferred policies may be drawn. An application of this method is made, to show its use and a comparison made with the utility approach. The final, and main, part of the paper examines the robustness of the method for both deterministic and probabilistic input criteria. In this examination it is seen that the mapping method performs very well in picking up extremes of preference even under severe tests of robustness.

THE BASIC algorithm in multi-dimension scaling, developed by Shepard and Kruskal [4, 7], takes as input the rank order required for the distances between the $\frac{n(n-1)}{2}$ pairs of a set of n points. Starting from an initial random configuration of these n points, the algorithm develops a mapping of them by successive iterations which attempts to produce as near as possible the required ranked ordering. The map can be made in any number of dimensions; if a K dimension map is required an initial fit is made in $(K+1)$ dimensions. This fit is projected parallel to the weakest of the $(K+1)$ principal components and further iterations made in the K dimensions. The goodness of fit of the solution is measured by the stress, S, defined as

$$S = \left(\sum_{i \neq j} (d_{ij} - \hat{d}_{ij})^2 \Big/ \sum_{i \neq j} (d_{ij} - \bar{d})^2 \right)^{\frac{1}{2}}$$

where d_{ij} = distance between points i and j;
$\bar{d}$ = average of d_{ij};
$\hat{d}_{ij}$ = numerical value chosen as close to its counterpart d_{ij} as possible, subject to being monotonically related with the pre-assigned dissimilarity measure δ_{ij}.

David Kendall developed this computer algorithm [1, 2] and applied it to points linked by other concepts of distance. Some of the most interesting applications have been in the seriation of graves from archaeological evidence, in mapping the position of parishes in the middle ages from ranked intermarriage data of that period, and in building maps of land ownership and use in rural settlements, also for that period from fragmentary data.

As an illustration of the method Kendall has shown [3] how limited information (for example, which pairs of the departments of France abut), can be used to draw a map of France in which the departments are correctly located relative to each other. In other words, strictly local information can be combined to produce global relationships. Other work [6] has used the abuttal concept to draw maps of organisational policies.

Let us suppose we have a set of alternative policies, each of which will result in different degrees of attainment of a range of incommensurable objectives. Most approaches to the problem of selecting an 'optimal' policy from such a set, use utility and we may then regard the 'distance' apart of a pair of policies as the difference between their two utilities. However, in order to obtain utilities, strong requirements are necessary—it is essential not only to be able to state preferences between any two policies, but also the strength of this preference as deduced from hypothetical gambles (for a discussion of the utility axioms from a behavioural viewpoint see [5]).

Vol. 5
1977

A much weaker requirement would be merely to ask for each policy in turn, what other policies give results which, taken as a whole, are roughly equally attractive. (Two policies assessed in this way as being equally attractive are called 'linked'.) It is most important to note that two policies can have results, in terms of their degrees of attainments of the range of objectives, which are quite different, while being similar in attractiveness. Such information is rather like the abuttal data of the French map. (It is also consequentially essential, that the transitivity axiom $A \sim B$, $B \sim C$; $\Rightarrow A \sim C$ need not apply.) Applying the algorithm, with Kendall's modifications, has shown [6] that it is possible to map policies in this way so that a small group of policies, at each end of the map, correspond to the most preferred and least preferred policies. That is, it is possible to deduce extremes of preference from indifference. This, if it is justified, is of some importance.

The purpose of this paper is to present, albeit briefly, a second application of this method to policy selection, to examine the relationship of the solution to a comparable utility solution and finally, to generalise and to ask how such a mapping would perform under tests of robustness.

An example of the mapping approach

Reference [6] gives an example of the approach by taking a hypothetical town of Brove in which there are 24 different ways of spending a fixed sum of

money to achieve five incommensurable objectives. It showed

(a) that consistent maps could be produced which throw 'best' and 'worst' policies to opposite ends;

(b) there was, in that case, a correlation, significant at the 5% level, between the position along the principal axis of the two dimension map and the corresponding utility, and

(c) there was evidence to suggest that a critical feature in the consistency of the map was that each policy should have at least two and preferably at least three links to other policies.

Further work (as yet unpublished) shows this mapping procedure to be robust under random link and policy omission.

We now give another (if rather trivial) example of the method. A note was made of 24 secondhand cars offered for sale in a local paper. The author assumed he could take any one of these as a gift and, whenever the price of a car was less than £1800 (the price of the most expensive) the residual amount was offered as a cash bonus. Hence a car with a price of £1400 would have a

TABLE 1. AN EXAMPLE OF ALTERNATIVE OFFERINGS

Cars on offer Reference No.	Year	Make	Mileage (if known)	Cash bonus	Indifferences
1	New	Mini	0	400	7, 14, 21, 24
2	1973	Mini	15,000	800	3, 5, 7, 10, 11, 20
3	1974	Datsun	10,000	400	2, 10, 12, 17, 18
4	1974	Cortina	30,000	300	6, 11, 14
5	1973	MGB Sports	39,000	400	2, 8, 9, 12, 14, 16, 19
6	1974	Renault 12L	32,000	400	4, 8, 11, 12, 14, 15, 17, 22
7	1976	VW Golf	7000	300	1, 2. 23, 24
8	1973	Triumph 2.5	?	0	5, 6, 16, 20
9	1974	Triumph Spitfire	16,000	300	5, 11, 18, 22
10	1972	Viva	36,000	1200	2, 3, 11, 13, 16
11	1971	Lancia 2000	?	800	2, 4, 6, 9, 10, 12, 13, 15, 22, 23
12	1973	Maxi	29,000	600	3, 5, 6, 11, 13, 17, 18, 20
13	1973	Hillman Hunter	?	1000	10, 11, 12, 16
14	1975	Chevette	?	100	1, 4, 5, 6, 17
15	1972	Triumph Spitfire	39,000	1000	6, 11, 22, 23
16	1972	Capri	?	600	5, 8, 10, 13, 21, 23
17	1971	Escort	?	1000	3, 6, 12, 14
18	1971	Mini	?	900	3, 9, 12
19	1975	Honda	?	1000	5, 20, 23
20	1974	Wolseley 6	?	300	2, 8, 12, 19, 24
21	1974	Mini	?	900	1, 16, 23
22	1967	Daimler	32,000	800	6, 9, 11, 15
23	1971	Capri	?	900	7, 11, 15, 16, 19, 21
24	1972	Granada	39,000	600	1, 7, 20

The table means that the author found equally attractive cars (1, 7), (4, 6), (5, 8), etc.

cash bonus of £400. We note that this makes the problem of selecting the preferred car more difficult since under free market conditions the 'average' purchaser might now be indifferent between all the cars.

The algorithm first of all evaluates the 'distances' between any two references in very simple terms—that is, adjacencies (directly linked policies), adjacent adjacencies and the remainder. It then attempts to find a location of reference points in two dimensions (it could be in any number of dimensions) so that the rank order of distances between the 276 pairs is as close as possible to the given order. A number of iterations were run each starting from a different random configuration. The results were very consistent and a typical solution is shown in Fig. 1. A short period of reflection by the author showed that he greatly preferred the extreme group (07, 01, 21, 24) to (09, 03, 18, 22) and hence the former is the preferred group.

The plot is drawn with the principal axis perpendicularly up the page. It was clearly of interest to enquire the relationship between perpendicular position and the corresponding utility. Hence the author took his separate utilities for make of car, year, and cash bonus, added them to produce the following:

Reference	1	2	3	4	5	6	7	8	9	10
Utility	2·1	2·25	1·45	1·35	1·75	1·75	2·18	1·45	1·85	1·45
Reference	11	12	13	14	15	16	17	18	19	20
Utility	1·55	1.85	1·72	1·45	2·12	1·85	1·32	1·90	2·12	2·05
Reference	21	22	23	24						
Utility	2·50	1·70	2·00	2·05						

Vol. 5
1977

The distance up the page was measured for each point and the corresponding correlation coefficient between this distance and the utility was calculated to be −0·56 which is significant at to 0·1% level, and is even more highly significant than in the Brove example. It is encouraging to note that there is such a close relationship.

The highest utilities in the above listing are for car references 21, 02, 07, 15, 19, 01 and the lowest are for 17, 04, 03, 08, 10, 14. Comparing these with the four car references at each end of the map we have (using the negative sign of the correlation coefficient):

High Utilities:	21,	02,	07,	15,	19,	01				
'High' end of map:	21,		07,			01,	24.			
Low Utilities:	17,	04,	03,	08,	10,	14				
'Low' end of map:			03,				09,	18,	22.	

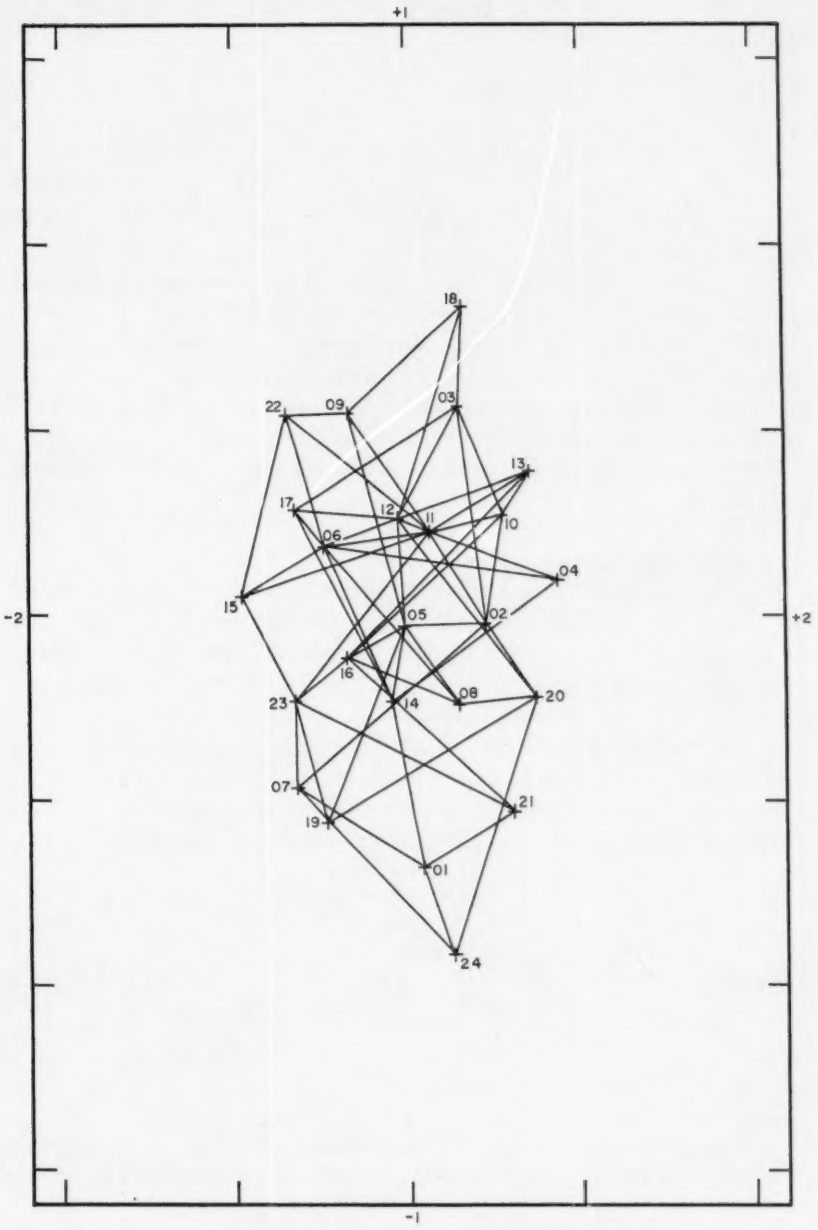


FIG. 1. *Map for* 50 *iterations* (*stress* 25·2%).

The low end is less encouraging but we note the corresponding utilities of the car references of the mavericks as:

24	2·05
09	1·85
18	1·90
22	1·70

and at least we can be pleased that the low end mavericks all have utilities less than the high end maverick (24).

One would not expect that the solutions of the two methods would be identical, indeed if the mapping solution gave an answer identical to that yielded by an application of utility, then some no doubt would use this as a valid objection to mapping *per se*. The point is that there does seem some relationship (which one ought to expect) but that the result is obtained with much less axiomatic demand than that made by utility theory.

The results so far therefore confirm the tentative conclusions of an earlier study [6].

An experiment in robustness and consistency

There is a clear difference between Kendall's maps and these. In Kendall's case there is a true and consistent solution in existence, completely known as in the districts of France, partly known for the historical case. Here however we do not know the true answer nor do we know the extent of the mutual consistency of the links adduced by the analyst. It seemed prudent therefore to invert the problem in order to test the validity of mapping answers under conditions of uncertainty.

We took a series of 25 two digit random numbers in the range 00–77. If these are regarded as the 'values' (v_i) of 25 policies we can order them as follows:

Policy (i)	1	2	3	4	5	6	7	8	9	10
Value (v_i)	05	07	09	12	13	17	19	20	22	26
Policy (i)	11	12	13	14	15	16	17	18	19	20
Value (v_i)	30	31	32	38	40	42	46	50	54	60
Policy (i)	21	22	23	24	25					
Value (v_i)	61	66	72	75	77					

We then assigned links between pairs of policies according to a considerable number of different criteria. For reasons of space we give here only five representative maps (the other results are obtainable from the author).

Criterion 1 (Deterministic). Two policies i and j, are assigned a link if

$$|v_i - v_j| \leq 10$$

Criterion 2 (Probabilistic). Two policies are assigned a link probabilistically depending on $|v_i - v_j|$: . . . .

$\lvert v_i - v_j \rvert$	0, 1 or 2	3	4	5	6	7	8	9	10	≥ 11
Probability of a link	1·00	0·80	0·65	0·50	0·50	0·40	0·35	0·20	0·10	0·00

Criterion 3 (Probabilistic).

$\lvert v_i - v_j \rvert$	≤ 10	>10
Probability of a link	0·50	0·00

Criterion 4 (Probabilistic).

$\lvert v_i - v_j \rvert$ (inc)	0–5	6–10	11–15	16–20	21–25	>25
Probability of a link	1·0	0·7	0·5	0·3	0·1	0·00

Criterion 5 (Probabilistic).

$\lvert v_i - v_j \rvert$	≤ 25	>25
Probability of a link	0·50	0·00

Such criteria seem to be in increasing order of their likelihood of introducing internal contradictions in the linkages.

It is interesting to note that Fig. 2 has a slight HORSHU effect [2] which can occur when the number of links per policy is small and this effect is much more pronounced in Fig. 3. In such cases the line of seriation is to be taken along the pattern of the 'horseshoe' and not up the page. As the possibilities of contradiction increase, so the diagrams get fatter, or to be more precise the second principal component becomes more important relative to the first. By the time we have progressed to the final Fig. 6 with its consequential internal contradictions, we are approaching the length–width ratio of the car example of Fig. 1.

However an examination of the extremes, which it is the purpose of the mapping method to expose, shows:

Figure	2	3	4	5
Four extreme 'maxima'	22,23,24,25	24,25,23,22,	25,23,24,22,	25,23,22,20
Four extreme 'minima'	01,02,03,04,	01,02,03,05,	02,01,04,05,	01,02,03,04,

	6
Four extreme 'maxima'	25,24,23,19
Four extreme 'minima'	01,05,03,02

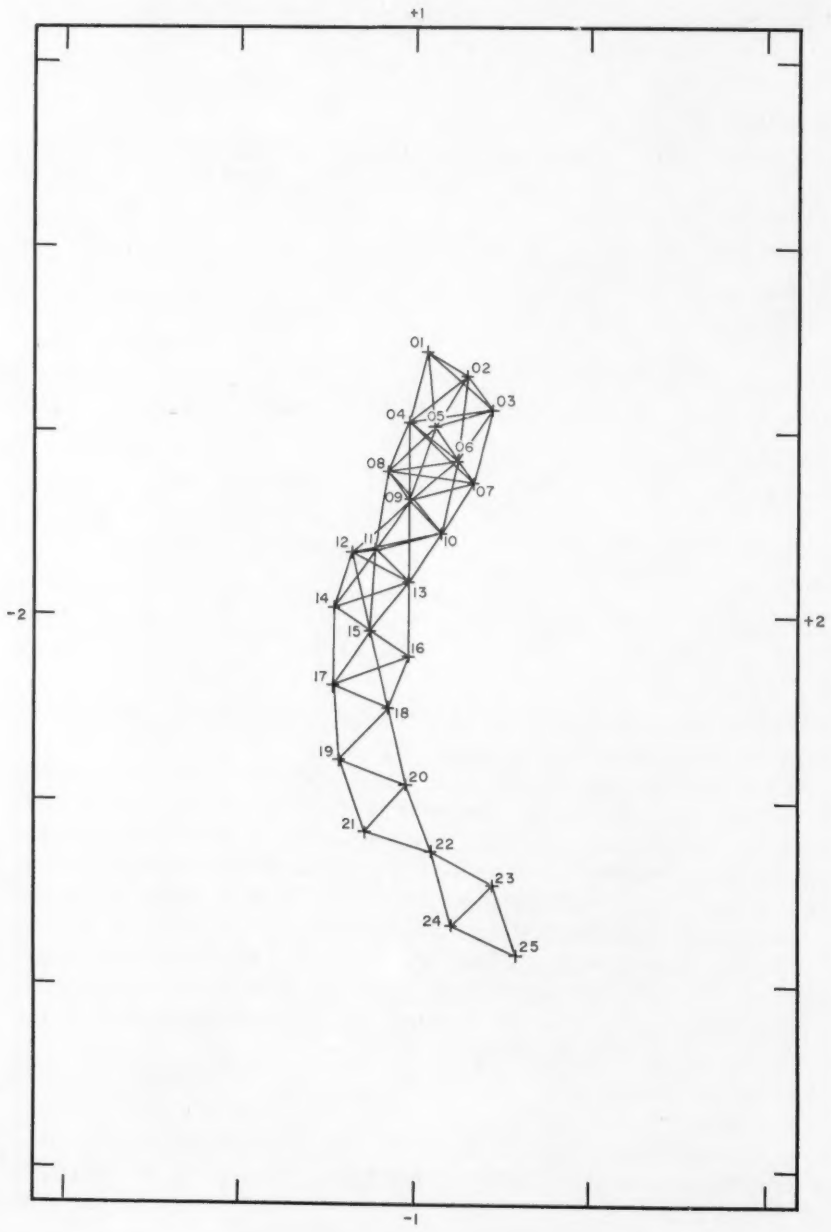


FIG. 2. *Map for criterion* 1 (*deterministic*).

Vol. 5
1977

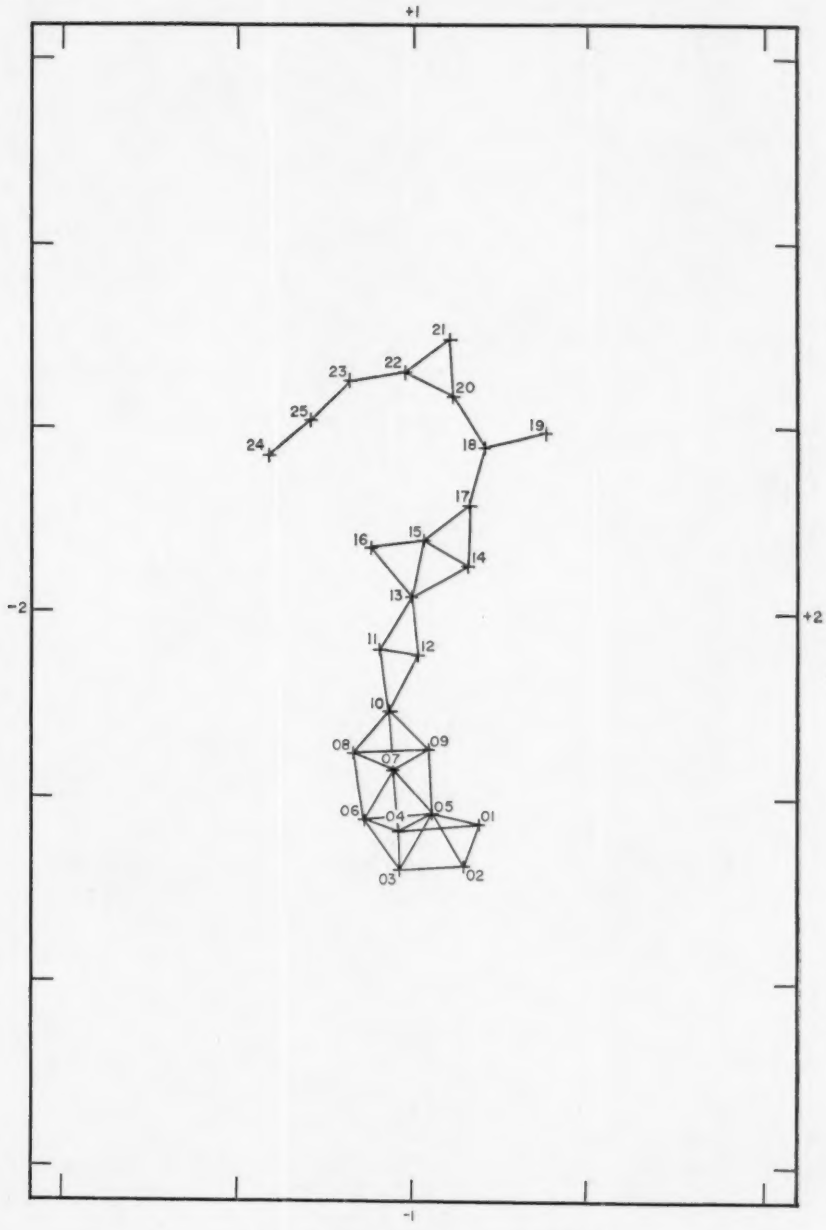


FIG. 3. *Map for criterion* 2 (*probabilistic*).

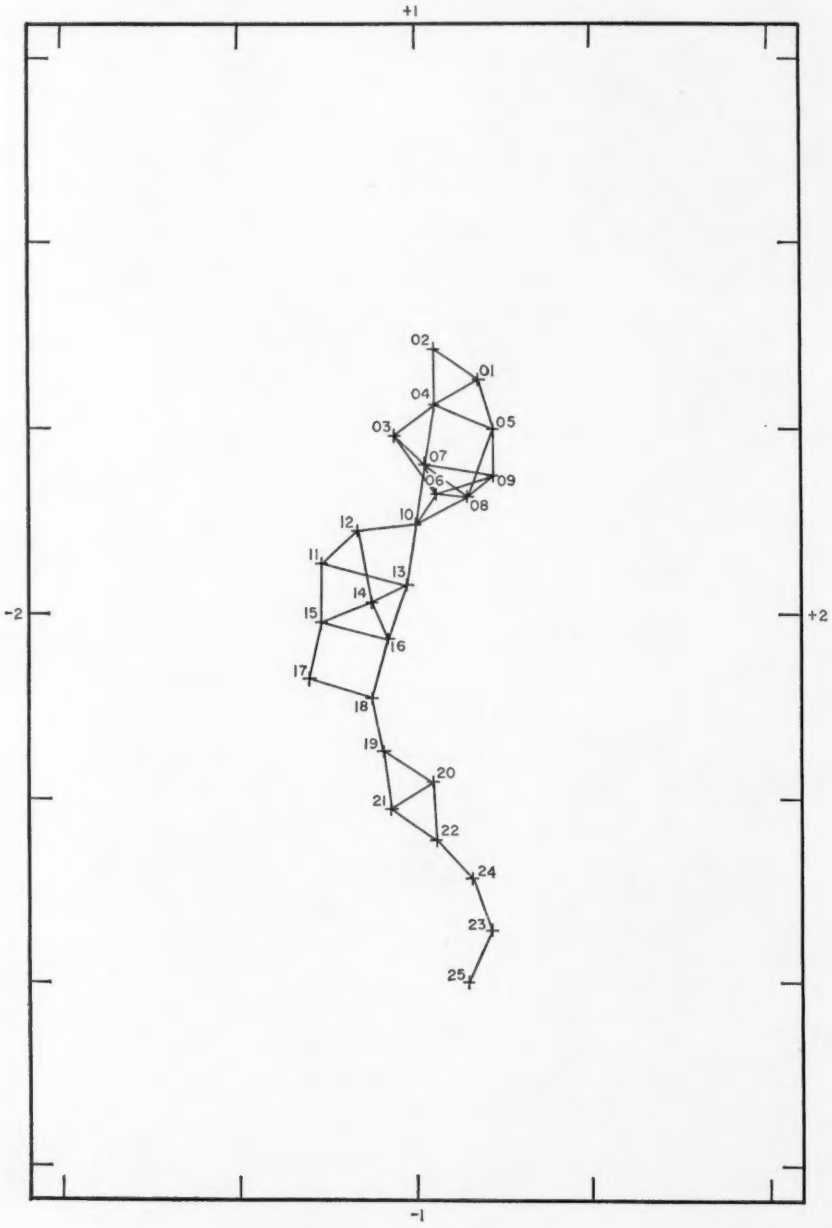


FIG. 4. *Map for criterion* 3 (*probabilistic*).

Vol. 5
1977

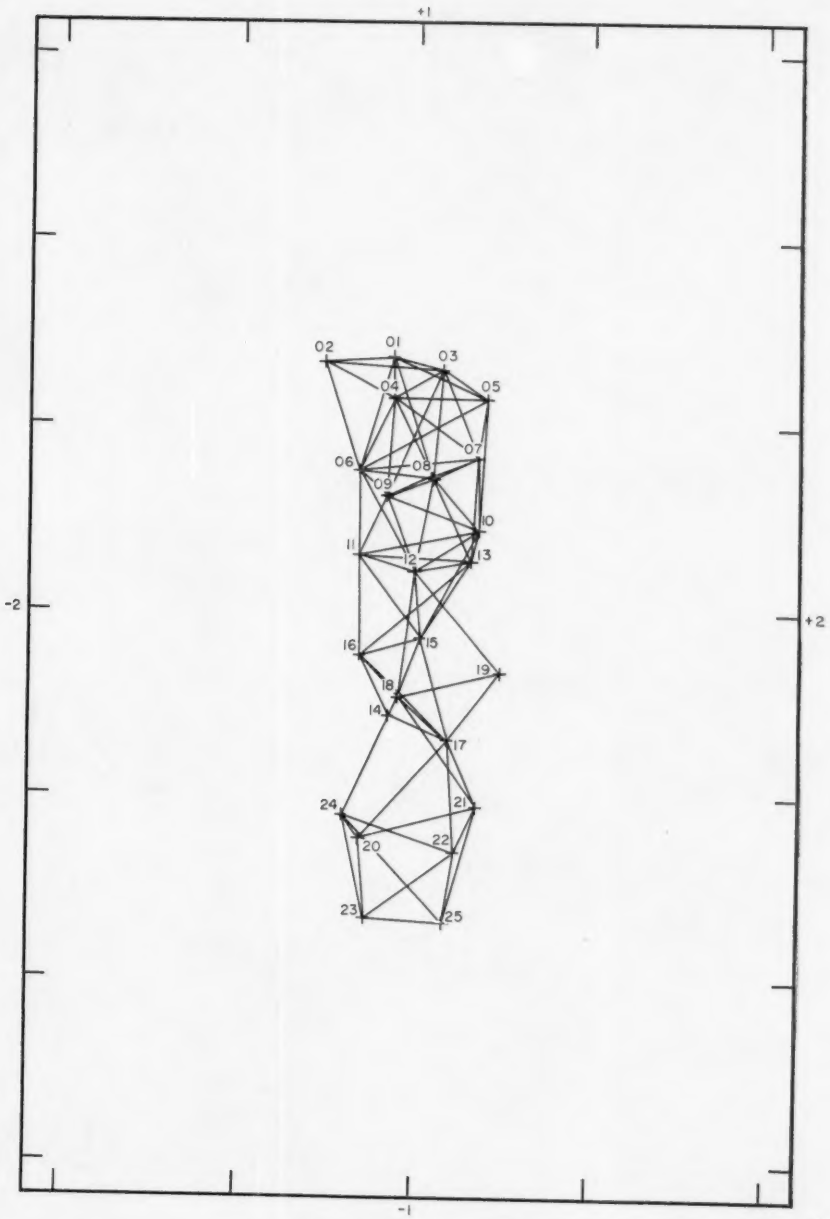


FIG. 5. *Map for criterion* 4 (*probabilistic*).

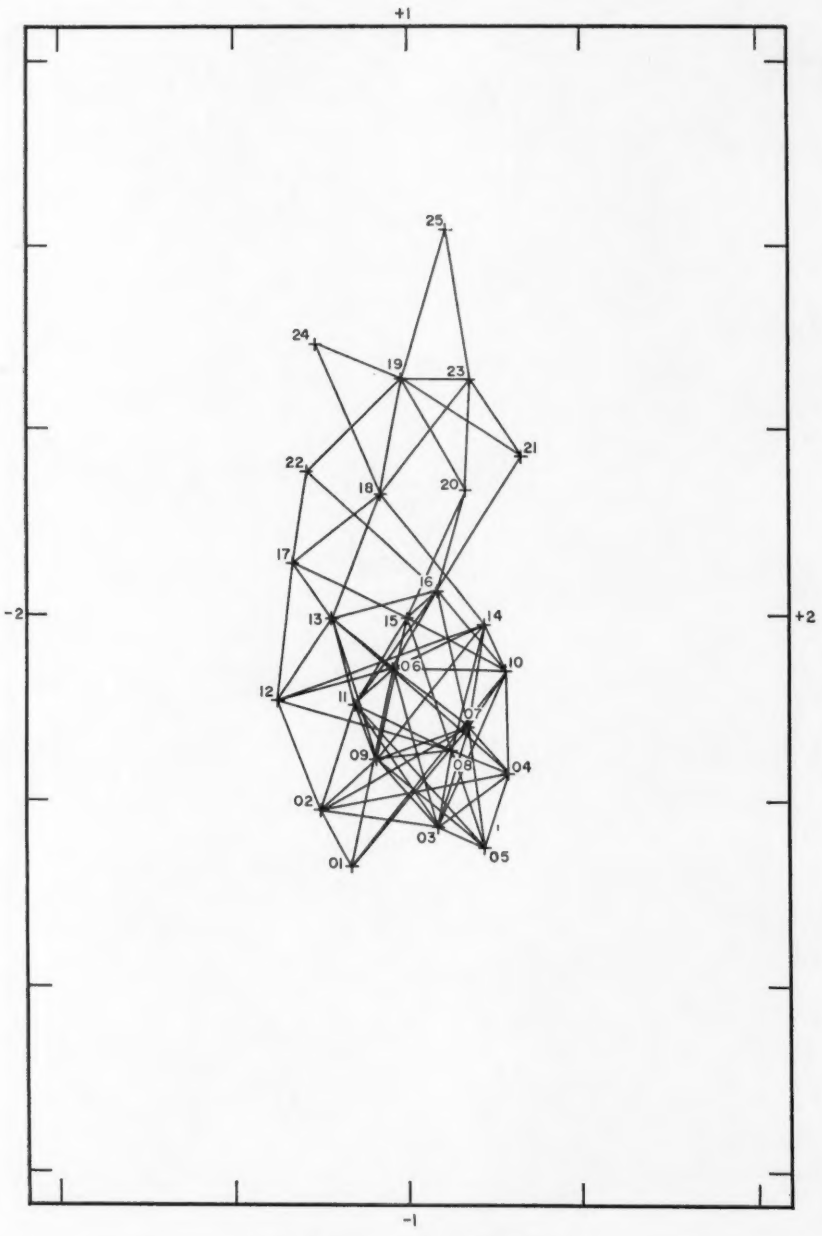


FIG. 6. *Map for criterion 5 (probabilistic).*

Vol. 5
1977

That is, 25 and 23 are omnipresent, 22 and 24 are present in four out of five. At the other end 01 and 02 are omnipresent, 03 being present in four out of five. It is indeed encouraging to note that even the extreme treatment of criteria 4 and 5 (Figs. 5 and 6) is not sufficient to seriously affect the robustness of the maps—the known extremes of 25, 24 and 23 and 01, 02, 03 being almost completely present.

CONCLUSIONS

Clearly no formal conclusions can be formally proved. However, there does seem evidence to suggest that even in an extreme case where the discrimination of perceived similarities is as fuzzy as in Criterion 5 (probability of a link of 0.5 for a difference as large as a third of the whole range) the algorithm applied in this new field to a purely binary input does yield results which in some sense are 'true'. In addition there appears evidence for the suggestion that the width of a map compared with its length is an indication of the internal consistency of the statements with which the algorithm is fed.

Finally, as was noted elsewhere [6], there is evidence that the points mapped on the perimeter tend to be those which have fewer links. For example, in Fig. 6, the 11 perimeter points 01, 02, 12, 17, 22, 24, 25, 21, 10, 04, 05, have a total of 53 links while the remaining 14 internal points have a total of 111 links, a result which in itself is of interest.

1. 5
77

ACKNOWLEDGEMENTS

I would like to thank Professor David Kendall of Cambridge University for his constant advice and encouragement, and also David Clarke of Sussex University, who carried out much of the analysis and computer runs.

REFERENCES

1. HODSON, FR, KENDALL, DG & TAUTU P (1971) *Mathematics in the Archaeological and Historical Sciences*. Edinburgh University Press, Edinburgh.
2. KENDALL DG (1975) The recovery of structure from fragmentary information. *Phil. Trans. R. Soc. Lond. Ser. A*, **279**, 547–582.
3. KENDALL DG (1971) Construction of map from 'odd bits of information'. *Nature* **231**, 158–159.
4. KRUSKAL JB (1964) Multidimensional scaling. *Psychometrika* **29**, 1–27, 115–129.
5. RIVETT BHP (1972) *Principles of Model Building*. Wiley, New York.
6. RIVETT BHP Policy selection by structural mapping. *Proc. R. Soc.* in preparation.
7. SHEPARD RN, ROMNEY AK & NERLOVE SB (ed) (1972) *Multi Dimensional Scaling*, Vol. 1, pp. 8 & 9. Seminar Press, London.

ADDRESS FOR CORRESPONDENCE: *Professor BHP Rivett, The University of Sussex, Operational Research, Mantell Building, Falmer, Brighton BN1 9RF, UK.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

A Review of Inflation Accounting and its Economic Implications

RJ BETTS

Imperial College, London

(*Received March 1977*)

The purposes of this paper are firstly to identify the problems which inflation creates in the context of conventional accounting methods, secondly to consider the alternative forms of inflation accounting which have been proposed, thirdly to look in more detail at the particular method which is to be adopted in the UK, and finally to consider the effects which that method will have on company results and on broader economic issues.

l. 5
77

1. INTRODUCTION

DISSATISFACTION with currently used accounting systems has grown in recent years partly because of the accelerated pace of inflation since the end of the 1960's and partly from a growing resignation to the fact of continuing inflation. The latter in particular has led to a recognition of the need for a systematic and uniform adaptation of accounting practice rather than the ad hoc and piecemeal adjustments to particular distortions to which accountants have hitherto had recourse.

If an accommodation to inflation is needed, not only by accountants but in all spheres of economic life, then an appropriate general strategy for this would be to identify the distortions and/or inequities to which inflation gives rise and then to devise the means whereby these undesirable effects may be minimized without at the same time creating other evils such as at one level, an acceleration of the inflation itself, or at another level damage to valued professional standards of financial reporting.

This being so, an appropriate starting point in looking at the particular problem of company accounts would be to ask, "what are the distortions which inflation gives rise to in company accounts as they are currently formulated?" In fact, this will be familiar ground to most readers and only a brief resumé of the main problems is required.

2. INFLATION AND THE HISTORIC COST PRINCIPLE

The long-established principles by which company performance and value are measured are derived from the historic cost principle.

Assets are valued in the balance sheet by reference to their price at the time of purchase. Liabilities are shown at their monetary value. Profit for the year is equated with all gains arising during the year and the capital to be maintained before profit is distributed is the monetary value of the shareholders' interest at the beginning of the year. This profit is determined by reference to the historic cost incurred in generating revenue for the year in accordance with the 'accruals' concept whereby revenue and costs are recorded as they are earned and incurred, respectively, rather than at the time when they result in money income or outgoings.

The fact that these principles have been accepted for so long would in itself suggest that they have much to commend them and that they should not be lightly cast aside. The great virtues which may justly be claimed are those of objectivity and verifiability. Any efficient system of resource allocation requires valid data. Thus participants in the capital market are bound to value highly any system of reporting which appears to minimize dependence on somebody's subjective judgement.

This is not to say that historic cost accounting is entirely objective. The basis of depreciation rates may be arbitrarily chosen, and other subjective judgements must be made in the assessment of bad debt provisions, the valuation of goodwill, the valuation of assets in the event of mergers and acquisitions, the treatment of research and development expenditures, and the conversion of overseas currencies in consolidated accounts. We might add that the conventions also call for the writing down of assets whose 'value to the business' has fallen below the book value derived from historic cost. Whilst subjectivity is inevitable in all these instances, they do not in themselves make a case for throwing out the historic cost principle.

The case against historic cost has to be made in the light of more general criteria against which any accounting system would have to be measured. The 'Sandilands Report' [11] offers the following list of requirements:

(a) The unit of measurement used in the accounts should meet a number of criteria relating to comparability, constancy, and use as a medium of exchange.
(b) The measurement of net assets should indicate their 'value to the business'.
(c) The concept of profit used should be capable of indicating the quality of the company's operating performance.
(d) The company's liquidity position should be readily discernible from the accounts.

In the light of these requirements the problems which arise with the unmodified use of historic cost in a time of inflation are not difficult to identify. Changes in the value of the unit of measurement (i.e. money) are not recognized in an appropriate way, the book value of the assets will only coincidentally represent their value to the business, the profit measure compounds profit deriving from operating performance with gains arising from the holding of appreciating assets, and liquidity is not discernible from the surpluses gained over the historic cost of assets consumed where these are less than the cost of maintaining assets for the purpose of continuing the business. In practice historic costs have had to be supplemented with statements to show the company's liquidity position.

The practical consequences of these are that:

(a) Balance sheets do not reveal the true value of assets except where they have been recently revalued for this purpose. Since revaluation has taken place in some cases but not in others, the principle of comparability is unsatisfied.
(b) Depreciation is inadequate to replace the value of assets consumed during the year.
(c) The charge shown for the cost of stock consumed is not sufficient to provide for replacement.
(d) The effects of holding monetary assets or liabilities are not shown [6].

In terms of the measurement of performance, growth is overstated since no adjustment is made for the fall in the value of the money unit in which results are expressed. Furthermore, whilst capital may be maintained in money terms, in real terms it may be distributed to shareholders and the Inland Revenue as well as to employees and customers by way of respectively 'excessive' wages and artificially low prices. Clearly such distributions can only be to the detriment of the long term viability of the company.

The deficiencies have become sufficiently serious for most commentators to accept the need for an alternative system of accounting which can meet more nearly, in a period of inflation, the requirements of a good inflation system such as those set out above. It is worth noting, however, at least in passing that at least one writer sees the pressure for change as resulting from confusion or even from the desire to confuse [9]. The writer argues that accounts are not valuation statements, that accounting figures are symbols of the existence of assets and not of their value. The purpose of the accounting figures is auditing and they should permit the auditor to account for all the firm's resources and to satisfy himself that they have been used in a proper way. Whether they have been used in a wise way is a separate matter. The appropriate measure of wisdom of use is return on capital and none of the proposed systems of inflation accounting, it is argued, take us any closer to that.

This seems to be an unrealistically narrow view ignoring, as it does, the uses to which accounts are actually put in which the figures are, necessarily for want of better information, taken as indicators of value.

Two major alternative approaches to accounting for inflation have been put forward, together with a third which incorporates some of the features of each of the other two. The main features of these methods will now be briefly examined before a more detailed treatment of the method which is intended for implementation in Britain.

3. CURRENT PURCHASING POWER (CPP)

This method, also known as 'General Price Level Accounting' [12] proposes to change only the unit of measurement rather than the principles underlying historic cost accounts.

To do this a 'purchasing power unit' would replace the monetary unit so that the balance sheet would then show the amount of 'purchasing power' contained in the company's net assets (as opposed to the amount of money), whilst the profit and loss account would show the company's gain or loss of purchasing power during the year. This would be brought about by applying a general price index to the figures in the conventional accounts thereby converting them from money to real (i.e. purchasing power) values.

The accounting profession has in the past tended to favour this approach. The Accounting Standards Steering Committee issued a Provisional Statement in 1974 [2] which suggested that a supplementary statement showing CPP adjustments be appended to historic cost accounts.

The virtues perceived in the CPP approach are firstly that it is relatively simple and requires the least disruption to an already familiar system (i.e. historic cost). Secondly, since a general price index would be used it retains some of the features of apparent objectivity associated with historic cost, i.e. the cost of assets purchased on any given date would be revised upwards by a similar proportion. Thirdly, and perhaps most importantly, it is a true system of *inflation* accounting in that it deals with changes in the value of money *per se*. It would thus be applicable equally to monetary assets and liabilities, to shareholders' equity and to the value of physical assets.

Whilst acknowledging the force of these arguments, Sandilands rejected the CPP approach. The Report argues that the users of accounts are not involved in a problem of general inflation but in changes in the value of particular assets owned by the company in question and with the current cost to the company of earning its revenue. Particular assets can and do change in value by much more or much less than the change in the general price index over the period. The Report even goes so far as to argue that inflation for the shareholder, in relation to the income he draws from the company, is not to be measured by a general price index since he does not buy unchanging 'baskets of goods' over

time. Needless to say, general price indices have been considered useful for a wide range of other economic purposes and, on this point, Sandilands has been sharply criticised. One writer [8] has turned the argument against the Sandilands Report by pointing out that companies themselves do not buy identical 'baskets' of assets either but would be influenced by the changing relative prices of assets over time. The use of special rather general indices would seem to imply that they do not respond to changes in relative prices. The validity of this argument really depends on how committed the company is to its current line of business and the amount of monopoly power it has in its markets.

The significance of these points will probably be clearer after consideration of the second major variation of inflation accounting.

4. CURRENT COST ACCOUNTING (CCA)

This approach was favoured by Sandilands and in a somewhat modified form by the Inflation Accounting Steering Group of the Accounting Standards Committee under the chairmanship of Douglas Morpeth [6].

Its supposed virtues may already be implied in part from the criticisms levelled against CPP. The most general of these is that it provides a measure of corporate value and performance of an individual and specific character.

The value of assets is to be their 'value to the business'. This is their 'deprival value', or the loss that the company would suffer if by some misfortune or mismanagement it were to be deprived of the services of the asset.

The concept of profit to be used would distinguish operating profit on the one hand from holding and extra-ordinary gains on the other. The latter category arises from the rising value of the company's stocks and fixed assets. The purposes of this dichotomy (for Sandilands) are firstly to distinguish the company's commercial success as an operating company from mere gains arising from increases in asset values due to inflation, and secondly to distinguish earnings available for distribution from sums which must be set aside for the purpose of providing capital for the maintenance of the physical capacity of the business. Profit calculated on a historic cost basis, by contrast, compounds operating profit and holding gains arising as the difference between the historic cost of assets and their value to the business. Under historic cost conventions, such holding gains would be available for distribution.

It should be noted at this point that the CCA proposals as presented by Sandilands do not provide for the revision of the value of monetary assets and liabilities or of shareholders' equity. This has been a major point of controversy over CCA and the main arguments on either side are reviewed below.

Sandilands' dichotomy between operating profit and holding gains reflects an attachment to a concept of profit which derives from JR Hicks' concept of income. This sees profit for the year as "the maximum value which the company can distribute during the year, and still be as well off at the end of the year as it

was at the beginning". As might be expected, the concept of 'well-offness' is not without difficulty and in many respects the controversy over inflation accounting pivots on precisely this point.

The measurement of assets as their 'value to the business' has also caused difficulty. The 'deprival value' concept is attributed to JC Bonbright. Sandilands thought that 'value to the business' would normally be replacement cost since were the company to be deprived of an asset it could be replaced, at some cost, by an identical asset. Only where replacement cost is greater than economic value (i.e. the present worth of the loss the company would suffer were it not to have the productive services of such an asset), would replacement cost not be 'value to the business'. If so, this is very fortunate since, as is well-known, many different economic values can be made to dance on the point of the economist's pen. However, in times of inflationary depression the number of instances of replacement cost exceeding economic values may be greater than supposed especially if CCA reduces the stream of net cash flows (as perceived by the valuer) to be derived from the asset.

Sandilands proposed that the replacement cost of assets should be derived from officially published indices of asset prices grouped by industry. The practical and conceptual problems inherent in such indices, viz. the probability of delays in publication and of matching particular companies to industry categories, led Morpeth to recommend the use of suppliers' quotations. In doing so, fresh problems of verifiability are likely to be introduced and the final word on this has probably not been heard.

5. SPECIFIC AND GENERAL PRICE-LEVEL ACCOUNTING

This method has also been called Current Replacement Cost, although there are several variations of the central idea.

The central feature is that specific replacement valuations as in CCA would be combined with adjustments using a general index as in CPP.

One principal purpose is to extend CCA to cover the problems which arise with respect to assets and liabilities expressed in money terms. Since the value of money is declining through time, it is suggested that holders of net monetary liabilities are making a real gain and holders of net monetary assets are making a corresponding loss. This issue is explored in more detail below.

A second principal purpose of this approach is to distinguish 'real' holding gains which arise from the appreciation of assets by more than the general index, from nominal gains in line with the general index. Nominal gains would be separated from income whilst real holding gains would not.

The argument for such a distinction derives from what is seen as the artificiality of the distinction between the use of gains for the replacement of

assets and the use of gains to invest in different assets [3]. There is no reason why the former should be favoured against the latter.

This separation of nominal from real gains would then require the use of two indices; a special index corresponding to that suggested by Sandilands and an index of general price movements to measure changes in the value of money.

Here again, the significance of the distinction between real and nominal holding gains would depend on the degree of the company's commitment to its current line of business. The fact that its replacement costs had risen more than the general index would not be significant to a company which proposed to continue in its present business regardless of such differential changes. The argument would then be, we must presume, that there is no reason why companies choosing to do so should receive favourable tax treatment by being allowed to 'understate' their profits (i.e. by excluding real holding gains).

6. AMPLIFICATIONS OF SOME ASPECTS OF CCA

The Morpeth proposals in some instances modified those of Sandilands and in others filled in the practical details of the latter's general principles.

One important modification involves what Sandilands called the 'Statement of Total Gains'. Total gains earned in a year is the difference between the company's net assets at the beginning and end of the year. This can arise either from value added by productive activity such that the amounts received or receivable exceed the value of assets consumed, or from increases in the value to the business of the company's net assets as a result of changing prices. The latter are called 'holding gains' or 'revaluation surpluses' and under the Morpeth proposals they are credited to a Revaluation Reserve Account. An important point here is that the revaluation surplus is not necessarily the amount which should be set aside for the continuing needs of the business, and this links up with the imprecise definition of the profit concept used in CCA. The revaluation surplus for the year and the current cost operating profit are brought together into an 'Appropriation Account' and there added to any appropriation to the Revaluation Reserve, and the dividend. The balance of the Appropriation Account is a transfer to, or from, general reserves.

Thus the question of what should be distributable and what should be retained is left to the discretion of the directors. Certainly current cost profit, whilst a relevant figure, is not identifiable with what is available for distribution. On the face of it this would seem to be at odds with the definition of profit proposed by Sandilands. More positively, even an imprecise definition of profit along such lines has the virtue that it relates more satisfactorily to the value of resources used up in earning the corresponding revenue than a measure derived from historic cost and we should also wish to distinguish 'profits' from 'liquidity'.

A second controversial aspect of CCA is, as suggested earlier, the treatment of monetary assets and liabilities. CCA in its Sandilands form does not provide for adjustment of these in line with inflation. The CPP method would show a gain (or loss) on monetary items in terms of its purchasing power units but it is still not clear whether or not such gains should be regarded as profit or whether or not such gains should be regarded as distributable.

A historic cost approach, using CPP only as a supplementary statement, would imply that such gains are part of profit since the opening shareholders' interest would not be reduced (in terms of purchasing power units) if the gain were distributed. However, on a different view of profit such gains would not be part of profit since to maintain the purchasing power content of the company's assets would require refinancing at a correspondingly higher money value than the old liability.

Gains on monetary liabilities would only be distributable by way of borrowing [10]. The revaluation of non-monetary assets (but not of monetary assets and liabilities) would normally reduce the company's gearing. The company could then raise debt capital to restore its gearing and the increased liquidity might then make distribution possible.

Morpeth, whilst not accepting the general arguments for the adjustment of monetary items, acknowledged the significance of real changes in the value of monetary items for some purposes. In arriving at the amount available for distribution by way of the Appropriation Account, directors would be given discretion to make allowances for monetary items. Additionally, a statement showing the effects on shareholders' funds of changes in the value of money is to be attached to the accounts. Doubts have been expressed however as to the notice that outside users of accounts will take of this discretionary and supplementary introduction of monetary adjustments. Certainly they do not go nearly far enough to satisfy the advocates of CPP.

A further change introduced by the Morpeth proposals relates to cost of sales. Following Sandilands, these call for the cost of sales to include the value to the business, at the time of consumption, of stocks and work in progress. Significant exceptions to this rule relate to work in progress carried out as part of a contract (as, e.g. in construction) on the grounds that where work is for a specific customer, work in progress is comparable to a monetary asset, and to stocks which are bought by, e.g. a commodity dealer for speculative purposes only.

To the extent that such a change produces a higher cost of sales than under the commonly used FIFO valuation, companies' reported profits will decline.

Similarly, CCA will result in higher depreciation charges. The value to the business of fixed assets consumed will now be derived from (usually) replacement costs. Thus depreciation is calculated as a proportion of the current value to the business of the asset spread over its estimated remaining life. To the extent that asset prices are rising this will result in a higher depreciation charge than one derived from historic cost.

7. CCA AND COMPANY RESULTS

Since CCA charges against revenue in times of inflation costs which are greater than those which would arise under historic cost, the conversion from a historic cost to CCA basis of profit calculation will in most cases reduce reported profits substantially, and in some cases quite dramatically. One estimate [5] suggests that current cost profits for the year 1976–7 would be about half the level of conventional profits with higher depreciation and cost of sales accounting for the reduction.

Those companies worst affected are, of course, those with high stocks and high depreciation relativity to profits, and in particular those whose activities are largely in the UK where inflation has been higher than in most other countries. High ratios of stock or depreciation to profits might arise either from the nature of the business, e.g. where the rate of stock turn is low or in capital-intensive industries, or from the fact that profits are currently depressed for some reason.

Thus in the current year the same source estimates that CCA profits would be lower than conventional profits by percentages ranging from 20 for Contracting, to 40 for Office Equipment, 60 for Heavy Engineering, 80 for Shipping, up to 115 for Textiles.

Needless to say, the percentage reduction in changing to CCA would be lower for any company the higher is the absolute level of profit calculated on a conventional basis.

The UK clearing Banks are in a special position since the CCA method does not adjust the values of purely monetary items. As holders of net monetary assets, the banks will be shown to have higher profits than under CPP and will thus attract higher tax. In periods of price control they will also find themselves disproportionately restricted.

The clearers have no stocks but they do have some non-monetary assets which they will be able to depreciate on the same basis as any other company. This would have produced a fall in pre-tax profits for 1975–6 of 8 per cent compared with historic cost whereas CPP would have led to a 20 per cent fall [4]. The effect of the new system on individual banks will depend very much on the proportion and location of their overseas business.

According to Laing and Cruikshank [7], the UK clearing banks pre-tax profits in 1975 would have been (£m):

	Historic cost	Current cost	CPP
Barclays	142	130	42
Lloyds	90	87	72
Midlands	85	74	49
Nat. West.	108	91	18

The banks may reasonably claim unfair treatment under the CCA proposals. Businesses holding net monetary assets must, for reasons of prudence, hold a certain level of shareholders' equity in conjunction with any given level of monetary assets. With inflation, the money level of such assets will rise even if their real value does not and the equity base must be correspondingly expanded from profit retentions [10]. Thus some part of the apparent profit of such businesses should be seen as being equivalent to retention to meet the rising replacement cost of physical assets in other types of business. Where this is not recognized by the authorities, net monetary asset holders are in danger of being unfairly penalized by taxation and price controls.

8. GENERAL ECONOMIC EFFECTS

It is important to bear in mind that, 'taxation effects' apart, the move to CCA or to any other form of accounting will not in itself affect the underlying performance of the company. What is to be changed is the formal description of the company. Any economic consequences which follow from that change must therefore derive from its effect in changing the perception of the company held by its management, employees, shareholders, creditors, suppliers, etc. This is not to say that such changes in perception may not in principle be important; they may be very important, but seem unlikely to be so in this particular instance.

The main changes in the company's published accounts will, in times of inflation, be that the value of assets will in most cases rise and the operating profit will fall relatively to the profit calculated on the historic cost principle. The degree of these changes will depend on the circumstances of the company in ways already described.

Larger companies, or more particularly those with relatively sophisticated management accounting and information systems, do not base their decisions on published accounts but on data collected and analysed for internal purposes.

Such data are likely to incorporate already many of the modifications to historic data that CCA requires. To the extent that CCA represents the company more revealingly, published CCA accounts therefore make this more revealing information available to outside as well as to inside users. In companies where management practice is less sophisticated, the need to separate operating profit from revaluation surpluses (i.e. holding and extra-ordinary gains in the Sandilands terminology) may be a discipline which leads management to provide itself with more relevant information than it might otherwise have had. This may then affect the internal allocation of its resources and its distribution policy in so far as these are affected by the distinction between operating profit and revaluation surpluses. Where the distinction is significant, the recognition of it can do nothing but good from the point of view of the resource allocation system.

Probably more significant than the reactions of management are those of shareholders and creditors. To the extent that asset values are seen to rise under CCA, then the company's attractiveness to both should rise, except in the probably unlikely event of their reposing much faith in the return on net assets as a criterion. More significance is likely to be attached to the apparent reduction in operating profit.

The stock market will probably continue to judge equities mainly on P/E ratios and dividend yields and Phillips and Drew [5] consider it likely that P/E will be calculated on current cost profits after tax. Their calculations of P/E ratios on the FTA index for 1976 and 1977 show an increase from 6 to 13 and from 5 to 10 respectively on the change from historic cost to CCA.

However, the analysts suggest that the market is already discounting the effects of CCA. The published P/E ratio has fallen from about 17 in the middle of 1973, since when inflation has accelerated, attention has been concentrated on the inadequacy of historic cost accounts, and the inevitability of some form of inflation accounting has been recognized.

Even though the general level of equities seems to be discounting CCA, this does not seem to be true of relative movements of equity prices and P/E ratios when compared with the relative magnitudes of the effects of the change to CCA on the performance of different industrial sectors. This is attributed to uncertainty as to how the different sectors will be affected given the uncertainty over the precise form which the new method will take and over the appropriate modifications to historic cost data.

It is inevitable that such uncertainty will arise and persist until the market has learned to use CCA accounts in extracting the information it requires, even when the form of the accounts has been finalized. This is certainly an area in which perceptions, as such, are crucial and it seems inevitable that there will be at least an initial reaction against those capital intensive, slow stock-turn companies which seem likely to see the biggest fall in their CCA operating profit.

Other users of accounts are the employees, or more especially, the trade unions. If operating profits are to be lower under CCA, how will this affect union attitudes towards wage claims?

The TUC (Trades Union Congress in the UK) did not give evidence to the Sandilands Committee although Lord Feather, a former Secretary General of the TUC, was a Committee member. At least one prominent trades unionist [1] has declared that the unions are unhappy at the effects which lower apparent profits will have on collective wage bargaining.

The crucial question is whether union negotiators will regard revaluation surpluses as well as operating profit as sources for wage increases. Certainly if such surpluses are to be retained as being essential to the maintenance of the business, and that is the basis of CCA, then to suggest that union negotiators will insist that wage increases may be paid from the surpluses is to suggest that they will act in ways which will harm the employment prospects of their

members. All that can be safely said is that the trade unions, like other users of accounts, will take time to adjust to the new system.

One of the most common objections to the various suggested methods of 'living with inflation', such as general indexing, is that such an acquiesence is likely to make the condition worse. Similarly, the worry has arisen that CCA, or for that matter CPP, will make inflation more rapid. The general argument is that when companies discover that their costs are higher and their profits lower than they had realised, they will react by putting up their prices to restore their profit margins to what they had previously thought they were.

Here again, the use of some form of adjustment for inflation in published accounts does not in itself change the company's situation. Previous decisions on prices will in many cases be made already on the basis of inflation adjusted data, i.e. data which already recognises the rising value of assets consumed.

Where such data has not been used systematically then in many cases pressure on company liquidity would soon reveal the need for price adjustments as sales income became inadequate to finance the replacement of assets consumed. In such cases, where prices were too low to enable the business to maintain itself, the move to CCA would simply hasten the recognition of the need to raise prices. The remaining cases would be those where the company enjoys a position of great market strength which it exploits by charging prices based on historic cost but with margins great enough to maintain the business without an explicit recognition of the effects of inflation. In the absence of price control such companies may, if the market permits, wish to raise their prices still further on the introduction of CCA accounts. It may be thought, however, that such advantageously placed companies who are not also sophisticated enough to have recognised already the effects of inflation, would not be sufficiently numerous to have much effect on the general price level.

Vol. 5
1977

9. CONCLUSION

Although historic cost accounts have important advantages in the form of objectivity and verifiability, the distortions which they suffer in times of inflation justify their replacement by some form of 'inflation accounting'.

Of the two main methods, CPP and CCA, the latter is better adapted to the type of 'managerial capitalist' mixed economy now ruling in most western industrialised countries. In such an economy, the involvement of companies in particular markets and particular technologies, the rise of the institutional investor tied into large companies have weakened the market mechanism and made of inflation a specific and structural rather than a general and monetary problem. CCA recognises this much more than does CPP.

Accounts based on CCA provide information more relevant to the needs of all users of accounts than do those based on historic cost. The fact that such

increased relevance may be, at least initially, an embarrassment to some companies is not a reason for not using them.

The inadequate treatment of monetary items provided for even in the Morpeth proposals, is a serious deficiency and threatens to penalise financial institutions such as the clearing banks. It may be that this inequity will be at least diminished in practice but should be much more satisfactorily dealt with also in the formal provisions. The current debate will no doubt continue even after the introduction of CCA. It would be remarkable indeed if the proposed methods work with such perfection that subsequent modifications prove unnecessary.

REFERENCES

1. *Accountancy Age* (1975) **6**(45), 10–11.
2. ACCOUNTING STANDARDS STEERING COMMITTEE (1974) Provisional Statement of Standard Accounting Practice **7**.
3. BAXTER WT (1976) The Sandilands Report. *J. Bus. Fin. Acc.* **3**(1), 115–123.
4. CLARKE T (1975) Impact on the clearers. *Banker* **125**(596), 1190–1191.
5. GIBBS M *et al.* (PHILLIPS and DREW, Stockbrokers) (1976) E.D. 18—Morpeth's Proposals.
6. INFLATION ACCOUNTING STEERING GROUP (1976) *Guidance Manual on Current Cost Accounting*, para. 10.
7. LAING and CRUICKSHANK, Stockbrokers (4.12.1976), quoted in *The Economist*, pp. 109–114.
8. MYDDLETON DR (1976) Why Sandilands is unacceptable. *J. Bus. Fin. Acc.* **3**(1), 97–104.
9. RAYMAN T (1976) Sandilands statement. *Accountancy Age* **7**(16), 14–15.
10. ROSE H (1976) Sandilands—nearly true but not quite fair? *Banker* **126**(599), 25–29.
11. SANDILANDS FEP (Chairman) (1975) Inflation accounting. *Report of the Inflation Accounting Committee*. Cmnd. 6225 H.M.S.O.
12. VANCIL RF (1976) Inflation accounting—the great controversy. *Harv. Bus. Rev.* **55**(2), 58–67.

ADDRESS FOR CORRESPONDENCE: *RJ Betts Esq, Department of Management Science, Imperial College, Exhibition Road, London SW7 2BX*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

The Multilevel Approach to Organizational Issues of the Firm—A Critical Review[1]

RICHARD M BURTON

Duke University, North Carolina

BØRGE OBEL

Odense University, Denmark

(*Received August 1976 in revised form January 1977*)

1. 5
77

This paper critically reviews the progress and state of multilevel decomposition models for the firm and its management. The presentation introduces and summarizes the fundamental ideas and also gives concise statements concerning advanced developments. The pricing approach, and the budgeting approaches are the two approaches considered. For managerial planning, initial starting solutions, iteration processes, and stopping rules are important algorithmic and organizational considerations. Further, in decentralized organizations, the motivations for cheating in planning, incentives for innovation, and reactions to uncertainty require attention. For these approaches one concludes that further theoretical work designed especially for managerial relevance needs to be developed. Despite these deficiencies, the practitioner need not wait for all the answers as the currently employed methods of pricing and budgeting for planning suffer the same deficiencies, and yet lack the power to choose better plans as the multilevel decomposition approach offers.

I. INTRODUCTION

THIS PAPER critically reviews the main ideas and contributions of multilevel decomposition models for the theory and operations of the firm. We view the firm as an organization in the Barnard [6, p. 79] sense: 'A system of consciously coordinated activities of two or more persons'. The formal economic theory of the firm usually considers the firm as a single unity with possible multiple inputs and multiple outputs. With this approach, many of the interesting and important aspects of the internal design of the firm are missed. For the practical

[1] This research was partly supported by funds from a grant to Børge Obel from Statens Samfundsvidenskabelige Forskningsråd, Denmark, and was carried out while Børge Obel was Visiting Assistant Professor, Duke University.

manager, as well as the organization theorist, this is too limited [56, p. 16]. A more general view of the firm immediately creates issues in coordination, conflict, consistency of sub-goals, communications, and control.

The basic model is the decomposition model for mathematical programming which has interrelated decision-making subsystems arranged hierarchically. The focus is on a set of organizationally relevant issues for decomposition models. Those issues include: the basic coordinative mechanism (price or budget), the structure of the firm, the decision making level, the planning period, the motivation for cheating, the motivation for innovation, and the influence of uncertainty. These are the variables for the organizational design of the firm.

Recently, excellent reviews of the multilevel planning literature have appeared with varying orientations. Although related, none focuses directly upon the firm and its organizational design. Hurwicz [38] reviews the accomplishments by economists for the design of resource allocation mechanisms for the economy as a whole. Our treatment of cheating is related to Hurwicz's discussion of 'incentive compatibility'. Jennergren's [43] review of decentralization concepts discusses the contributions by organization theorists and accountants as well as economists. Ruefli [67], considered behavioral models for multilevel resource allocation. These models are particularly interesting as they permit the optimal solution to be a function of the organization structure. Freeland and Baker [24] extended Ruefli's work and introduced an entirely different coordination procedure to overcome the insufficiency of prices. Atkins [3] emphasized the managerial considerations of implementation for a decentralized firm. All of these reviews are relevant and related, but each adopts a different point of view than the one assumed here.

The following sections present a comprehensive statement concerning the theoretical contributions and practical implications for the multilevel theory of the firm.

Vol. 5
1977

II. THE PRICING APPROACH TO MULTI-LEVEL CO-ORDINATION

The basic idea of iterative decentralized planning using the pricing approach was developed by Dantzig and Wolfe [20, 21]. They developed a decomposition algorithm for solving large-scale linear programs but recognized that their approach could serve as a model for decentralized planning as well.

The organizational interpretation was treated in depth by Baumol and Fabian [7]. Their particular model is a two level structure with an executive (program) and two divisions. The mathematical model which represents the firm's decision problem was given as:

$$\begin{aligned}
\text{max:} \quad & P = P_1 X_1 + P_2 X_2 + P_3 Y_1 + P_4 Y_2 + P_5 Y_3 \\
\text{subject to:} \quad & a_{11} X_1 + a_{12} X_2 + a_{13} Y_1 + a_{14} Y_2 + a_{15} Y_3 \leqslant C_1 \\
& a_{21} X_1 + a_{22} X_2 \leqslant C_2 \\
& a_{31} X_1 + a_{32} X_2 \leqslant C_3 \\
& a_{43} Y_1 + a_{44} Y_2 + a_{45} Y_3 \leqslant C_4
\end{aligned}$$

where the P's are profit co-efficients, a_{ij}'s are input/output coefficients, C's are fixed resources, and X's represent division 1's decision variables with Y's for division 2. It is the special structure of this problem which lends itself to decomposition and the analogous hierarchical model of the firm. First, note the objective function (for profit) is linear, and therefore separable, and thus the portion with the X's can be assigned to division 1, and the Y's to division 2 without interference (or, externalities). Next, the first constraint affects all decision variables of the firm, and is called the corporate constraint. The next two constraints involve only the X's and the last constraint only the Y's. These are called divisional constraints. Similarly, C_1 is called a corporate resource, and C_2, C_3, C_4 are divisional resources.

The organizational algorithm consists of the executive, or coordination unit, solving a problem and then passing information to the divisions. The divisions, in turn, solve a planning problem and send information back to the executive. For a linear programming model, each submodel is also a linear program for the Dantzig–Wolfe approach terminating at an overall optimal solution. The process is given in more detail. Let us begin with any division's problem at an arbitrary iteration, e.g. division 1. Division 1 is asked to solve the following LP:

$$\begin{aligned}
\text{max:} \quad & R_1 X_1 + R_2 X_2 \\
\text{subject to:} \quad & a_{21} X_1 + a_{22} X_2 \leqslant C_2 \\
& a_{31} X_1 + a_{32} X_2 \leqslant C_3 \\
\text{all} \quad & X_i\text{'s} \geqslant 0
\end{aligned}$$

where R_1 and R_2 are modified profit co-efficients which are communicated to division 1 by the executive program. (We shall discuss their informational content later.) Division 1 then sends the optimal values of X_1, X_2 back to the executive unit. The executive program consists of the original objective function, the corporate constraints, plus a linear combination of all past solutions sent up by the divisions. (The details are given fully in [7].)
The executive program is:

$$\begin{aligned}
\text{max:} \quad & P = P_1 X_1^a + P_2 X_2^a + P_3 Y_1^a + P_4 Y_2^a + P_5 Y_3^a \\
\text{subject to:} \quad & a_{11} X_1^a + a_{12} X_2^a + a_{13} Y_1^a + a_{14} Y_2^a + a_{15} Y_3^a \leqslant C_1
\end{aligned}$$

$$\Sigma \mu_t = 1$$

$$\Sigma \nu_t = 1$$

all $\quad \mu_t, \nu_t \geqslant 0.$

The decision variables are the linear weightings, i.e. μ's and ν's. The X^a's and Y^a's are aggregated variables and are written fully as

$$X_i^a = \mu_1 X_i^1 + \ldots + \mu_t X_i^t \text{ and } Y_i^a = \nu_1 Y_i^1 + \ldots + \nu_t Y_i^t.$$

where the superscript t refers to the optimal solution submitted by the division at the t^{th} iteration.

Except for the last iteration and optimal solution, the executive primal program is not of great interest. The interesting information is in its dual. The dual variable for the corporate constraint, call it π_1 for C_1, measures the marginal value, an opportunity cost of that corporate resource. Divisions which use this resource are then 'charged' for its use by subtracting the 'cost' of the resource from their profit co-efficient. The modified profit co-efficient for X_1 is given as:

$$R_1 = P_1 - a_{11}\,\pi_1.$$

That is, the net profit co-efficient for X_1 is the original profit co-efficient minus the cost of the use of the scarce corporate resource. Similarly, modified profit co-efficients are calculated for X_2 and the Y's. Briefly, the executive sends modified prices (hence the name pricing approach) to the divisions, the divisions calculate their best solution for outputs, and send those back to the executive who in turn determines the value of the scarce corporate resources, and adjusts the prices. The executive solves for prices, the division's for outputs.

The pricing approach applies to non-linear models also. Malinvaud [51] presented a general formulation for concave/convex models with a slightly different model structure than the structure presented above.

In a vast literature about the pricing approach special versions have been treated. Different model structures have been presented by [10, 12, 14, 15, 29, 30, 31, 35, 36, 41, 75, 77]; different price adjusting techniques by [4, 9, 18, 63].

A general formulation of the Dantzig–Wolfe approach on concave/convex models without requirements of separability of the objective function or separability of the divisional constraints was given in [36]. If separability is not required, the co-ordinating unit must augment prices or profits, with information on the performance of other divisions to modify the objective function for the divisional programs, i.e. price function rather than a point value (e.g. [31]).

The initiation phase

In general, to start the planning cycle, the following issues have to be clarified: (1) which level in the organization initiates the planning procedure; (2) which starting strategy does the level use; and (3) does the initiation level have the information needed about the organization to initiate the planning efficiently.

Vol. 5
1977

The coordination unit may initiate the planning cycle by setting prices on the common resources and submitting this information to the divisions. If the coefficient matrix for the common resources has a special structure, then the procedure may be efficiently initiated by using prices close to the optimal dual variables [17, 49]. However, prices *equal* to the optimal dual variables are not a good starting strategy because they cause infeasibility of the coordinating program. The rules selecting the prices close to the optimal dual variables assume that the coordination unit has information about the coefficient matrix for the common resources, which in some organizations may not be available to that unit.

The starting strategy could also be based on prices from previous planning periods or if a market exists the market price on the common resources. These last strategies require no information about the corporate constraints.

If, on the other hand, the divisions initiate the planning cycle they must submit a planning proposal of their activities. The divisions may determine their own price forecast (based on the price used in previous planning periods) and solve their divisional program to obtain their planning proposal. Or, they could modify their plan from a previous period so that it still is feasible.

If the planning procedure is terminated before optimality is reached, e.g. according to a time constraint [54], the starting strategy has a great influence on the outcome of the plans [8, 17].

The iteration phase

Most of the literature on pricing methods has been addressed to the iteration phase; problems such as feasibility, the information flow, convergence rate, upper and lower bounds on the objective function, etc.

The two factors of interest in a planning context are the feasibility and the type of the information flow. There is no guarantee that the above starting strategies yield a feasible solution for the coordinating problem. However, it is very important both for a truncation and an interpretation of the intermediate planning proposals that the coordinating problem be feasible; see [49, 51].

With respect to the information flow a great many possibilities are present, e.g. see [2, 39]. An analysis of these possibilities in the linear case is given in [49]; it is shown that 12 different possibilities exist, each inducing different basic knowledge of the organization.

The termination and implementation phase

Two main questions arise in the termination and implementation phase, namely when to terminate, and who is going to make the final decision.

The first to address the truncation question was Marshak [54] where he restricted the planning to take place within a certain time span. This can be combined with a maximum number of iterations; economic measures could

also be taken into consideration. As an example, we mention the difference between the increase in the objective function by making one more iteration and the cost of making the iteration. Obviously, neither of these two figures are easy to calculate. But the increase in the objective function could be based on upper and lower bounds, which are easily calculated, and the cost could be based on certain standards.

The truncation also makes the feasibility of the co-ordinating problem very important. When the planning proceeds for only a few iterations, and not necessarily to an optimal solution, the actual plan to be implemented must be feasible.

The question about decision level was considered by Baumol and Fabian [7, p. 14]:

> "It is to be emphasized that now division managers must be told by the company what weights they are to employ, i.e. what combination of their proposals the company desires them to produce. There is no automatic motivation mechanism which will lead division managers to arrive at such a combination of outputs of their own volition. In this way, the decentralization permitted by decomposition breaks down completely at this point."

Vol. 5
1977

They (p. 15) continue by 'we may conjecture that this difficulty is, in part, a consequence of the linearity of our problem". Indeed, this is the case as others have shown. The basic idea is quite simple. Any given division can only select production values from its own feasible region as the modified prices change the divisional objective function—not the divisional feasible region. Yet (p. 16), "corporate constraints may limit the solutions to a subregion of the divisional feasible region". That is, the global optimal solution may lie interior to a divisional feasible region and there is no way to convey this information from the executive to the divisions within the strictly linear framework. Various solutions to this criticism have been offered—most involve some departure from the strictly linear framework.

Rigorous proofs that a separable strictly concave objective function (for maximization) is a sufficient condition to permit the divisions to calculate their own optimal solution for implementation were offered in [16 and 57]. As conjectured by Baumol and Fabian, it is the linearity of the model which causes the breakdown—not the use of pricing systems *per se*.

Jennergren [40, 42], in a different approach, has developed a scheme which permits the divisions to make their own calculations of production which are optimal for the firm as a whole. Jennergren's approach is to send a price schedule (or supply function) at the final iteration. The particular mathematical form for the price schedule is linear; thus each division is then faced with a quadratic objective function to optimize. Again, Jennergren's result demonstrates that it is the linearity assumption which destroys the price scheme, and

not the pricing approach itself. However, Jennergren's approach does not work in connection with truncation.

Another approach [39, 49] is to tell each division the allocation of the common resources (this information is available from the co-ordinating program) and make that a specific constraint in the divisional problem. Each division then has a linear program where the original profit coefficients appear in the objective function (modified co-efficients can lead to an incorrect decision). The interpretation of this phenomenon is that, once the allocation is determined and the price thereby fixed the division has to pay $\pi_j C_j^1$ (C_j^1 being the allocation) for its usage of the common resources. Therefore, from a divisional point of view, the payment for the common resources used in the implementation is a fixed cost and is not relevant to the decision. This is independent of the truncation method used. Charnes *et al.* [16], rather than making a formal inequality, put a high penalty on violating the constraints, which gives the same result. Kydland [48] has also dealt with the termination problem, using a procedure where a ranking of the divisions is needed.

Experimental results and model building

Even though the price-directive method is the most analyzed method, very few reports have appeared dealing with aspects of practical implementation, model building, and the implications for the organizational structure.

Ljung and Selmer [49] analyzed the interpretation of the Dantzig–Wolfe method on three Swedish firms. Obel and Christensen [61] made the interpretation for a Danish firm and a regional model for Danish Agriculture. Kutcher [47] and Ross [68] both used the Dantzig–Wolfe method to discuss the regional structure of Mexico and Ireland respectively.

Simulation experiments on small randomly generated models were described in [44 and 50]; experiments on bigger existing models were performed by [17, 49, 78]. The general results from these experiments are that the number of common constraints is critical and should be kept as low as possible, the number of divisions should not be 'too small', and an efficient starting strategy is very important. The organizational design implications of these remarks will be developed later in the paper.

As yet, no successful implementation of the general idea of iterative adjustment of the price approach for an existing firm, neither for the operational nor strategic planning, has been reported.

III. THE BUDGETING APPROACH TO MULTILEVEL CO-ORDINATION

The budgeting approach is the dual to a pricing approach. Basically, the executive, or central coordinator sends resource allocations to the divisions.

The divisions send prices (or marginal values) back to the central coordinator, who then adjusts the resource allocations until an acceptable allocation is determined. The underlying economic principle that adjustments should be made until a resource yields equal marginal return in all uses is well known. In terms of the multilevel approach to organizational issues, the budgeting approach has received less attention than the pricing approach. Yet, practical budgeting systems are widely used for planning and control in organizations.

One presentation of the multilevel budgeting approach for the firm is given in [13]; their paper gives a firm interpretation of the budgeting approach of [46]. The linear programming model presented at the beginning of the pricing approach section suffices to demonstrate the procedure. Division 1 would have the following problem:

Division 1

$$\text{max:} \qquad P^1 = P_1 X_1 + P_2 X_2$$

$$\text{subject to:} \qquad a_{11} X_1 + a_{12} X_2 \leqslant C_1^1$$

$$a_{21} X_1 + a_{22} X_2 \leqslant C_2$$

$$a_{31} X_1 + a_{32} X_2 \leqslant C_3$$

$$\text{all} \qquad X_i\text{'s} \geqslant 0.$$

where C_1^1 is the allocation by the central coordinator of the scarce corporate resource. The division problems are always of this form and they change for each iteration only by new allocations of C_1^1. Upon solving their problems, each division sends back to the central coordinator the marginal value of the corporate resources, i.e. the dual values of C_1^1 and C_1^2, respectively (C_1^2 being the allocation for division 2). Let those values be λ_1 and λ_2; if λ_1 is greater than λ_2 then the first division is making better use of the corporate resource than division 2 and hence division 1 should receive more and division 2 less. The formal linear program for the central co-ordinator suggests the results. The problem is:

$$\text{maximize:} \qquad \lambda_1 C_1^1 + \lambda_2 C_1^2$$

$$\text{subject to:} \qquad C_1^1 + C_1^2 \leqslant C_1$$

$$\text{all} \qquad C_1^i\text{'s} \geqslant 0.$$

Basically, the problem statement suggests that the central coordinator wants to maximize the value of the scarce corporate resource subject to the constraint that it can allocate no more than is available. The optimal solution to this LP is to assign all of C_1 to either C_1^1 or C_1^2 depending upon which of λ_1 or λ_2 is larger. It is noted that this solution must be tempered in the following manner. Not all of the scarce resource is allocated to the division with higher marginal return, but only a small increase over what it had on the previous iteration.

Vol. 5
1977

This tempering is needed to prevent unlimited oscillating allocations (or cycling) between the two divisions. Kornai and Liptak [46] use an arbitrary small increase (the change decreasing with each iteration) of resource to the division with the highest marginal return. The general economic principle of equal marginal return in all uses is a sufficient condition, but not necessary. Freeland and Moore [25] have shown that degeneracy in the divisional program is the usual situation, not the exception. They further show that equating of the marginal returns will seldom, if ever, occur at optimality. Thus, the optimality conditions for these algorithms are more involved than the usual marginal returns condition. Other budgeting approaches have been developed, all having the same divisional programs but differ in the way adjustments of the common resources are made, see [1, 26, 32, 33, 60, 71, 73, 79]. In [60] and [74], a resource-allocation approach is presented but the basic model is slightly different from the model presented above; their approaches are similar to [73].

The initiation phase

Contrary to the pricing approach, which is very flexible with respect to where information could be located before the planning is started, the location of the required initial information is fixed in the budgetary approach. For all the procedures mentioned, the headquarters must know the quantities of the shared resources available, i.e. the corporate right-hand side. The divisions must have information about the availability of divisional resources, and the objective function co-efficients for all potential divisions decision variables.

Starting strategies for these budgeting approaches is a relatively untouched topic. For the Kornai–Liptak approach, there is no research to suggest that an arbitrary allocation is inappropriate. The only required characteristic is that this allocation (and, all subsequent allocations) be feasible. It is interesting to note that if the initial allocation were optimal, the K–L approach may not detect it (if the divisional programs are degenerate) and further it may depart from the optimal allocation, only to return to it after a potentially infinite number of iterations. The Ten Kate [73] method is the dual to the Dantzig–Wolfe algorithm. Thus, starting strategy studies for D–W can be interpreted for the dual formulation; e.g. the heuristic rules for determining the initial prices can be applied to the dual.

If we turn to real world budgeting systems, we find an extreme reliance on last year's budget as the starting point [22]. As any budgeting system will be used in a dynamic situation, the answer to the starting problem may be suggested by the real world applications. However, no investigation of the efficiency of this dynamic approach has been reported. The research in selecting starting budgeting strategies is sparse and a general strategy for choosing the starting point remains an unresolved question.

The iteration phase

The budgeting approach has been investigated in terms of convergence, information flows, and feasibility. All methods iterate to an optimal solution. This may be little comfort in practice, as the number of iterations to a reasonable solution is operationally more important. Madsen [50] found the Kornai–Liptak method to approach an optimal solution only after some 40–50 iterations for small test problems. Variants on the Kornai–Liptak method were found to improve the situation, but the choice of good stopping rules is still not solved. Madsen found the Silverman method [72] to take very small steps and thus the earlier concern for a good starting strategy is evident.

Perhaps the more encouraging aspect of budgeting procedure deals with the information flow and feasibility. For the K–L method, assuming regularity of the associated game, any iteration is feasible. For Ten Kate, some intermediate solutions and the final optimal solution are feasible. For these feasible solutions, the divisions can calculate implementable solutions. However, some intermediate solutions may be infeasible. The Silverman method maintains feasibility but the required directional derivatives are complex to calculate, and each feasible step is quite small.

Information flows are smallest for the K–L approach where budgets are sent down and shadow prices are sent up. Ten Kate may also require divisional profits to be sent up; this is not a great burden. For the Silverman approach, the spanning vector for the directional derivatives may be a large data set to transmit to the central coordinator. All of these methods permit the divisions to make the final implemented solution, and further the central coordinator does not know the details of what will happen, i.e. it does not know the value of the X_i's or Y_i's for any division.

Vol. 5
1977

The termination and implementation phase

Stopping rules for budgeting approaches are not based on the economic principle of equal marginal return in all uses. This is a sufficient condition, not a necessary one, and will not be met in the degenerate case. For Kornai–Liptak and Ten Kate, their stopping rules take the form of differences between upper and lower bounds on the optimal solution as calculated at each feasible iteration. This difference provides an upper bound on the return from further calculations. Little work has been done to suggest an operationally appropriate value for this difference as a stopping point. For Silverman, the stopping rule is that no allocation changes from the previous iteration.

Experimental results and modelling

Very few reports are available concerning implementation and numerical aspects. Madsen [50] made a numerical study of several algorithms but the size of the test models was very small.

To the authors' knowledge, the only studies on real world data using a

resource-directive approach are reported in [17 and 61]. In these studies the Ten Kate procedure was used in connection with a slaughterhouse model and a model for Danish Agriculture. The coordinating mechanism in the first iterations made extreme allocations, but good results were obtained when the allocation was restricted by *a priori* determined bounds. These bounds were set by knowledge of previous performance of the organization.

The development of appropriate bounds may attract more attention in future research. One way to determine such bounds is to set a small range based on the allocation in the previous planning period. This method seems very attractive if the model does not change dramatically from one period to the next. Another way is to determine the lower bounds on the allocations in the current period. That is, if some products are sold before the planning period the lower bounds would be the amount needed to produce what has been presold. This is similar to zero-based budgeting.

IV. CHEATING, CONTROL, AND MOTIVATION

Cheating, as used here, is the conscious choice of a division unit to make decisions which enhance its own payoff (either absolutely or relatively) at the expense of the overall payoff and/or the payoff to other divisions. There is little doubt that this possibility exists for both the price directed schemes, and budget directed schemes. However, it is less certain whether it is in the interest of the division to cheat or not. It appears that it is more advantageous for small divisions to cheat than for large ones. The research on these issues is sparse, and there is opportunity for further contribution.

Hurwicz [38], discusses cheating under the topic of incentive compatibility and suggests the structure as a non-cooperative game without side payments. He states a general proposition (p. 24) that there exists:

> "A fundamental conflict among such mechanism attributes as the optimality of equilibria, incentive-compatibility of the rules, and the requirements of information decentralization. Concessions must be made in at least one of these directions."

Optimality of equilibria and information decentralization are compatible for calculating optimal *operating rules*. But incentive-compatible rules do not follow for price-takers [38]. Thus the *enforcement* issue is not solved automatically. Cheating is a real possibility for the multi-divisional firm. We review some of the work to determine just how large the problem is.

Jennergren [39] presented two models for cheating. One is a general model which builds upon the work of Marglin [53]. The other uses the Dantzig–Wolfe situation. The results for the D–W scheme are quite interesting. A division can enhance its own payoff by not revealing correct information concerning its

production possibilities. The scheme is as follows: the division reveals only a portion of its feasible production region. This reduces the demand for scarce corporate resources—which will be priced in equilibrium when supply equals demand. With the false information, the equilibrium price is set low, and the division can increase the payoff—namely, these resources are more valuable to the division than it reveals.

Jennergren's scheme for cheating is intriguing, but he argues that the scheme is not likely to be used as it requires the cheating division to have complete information concerning the central co-ordinator and other divisions as well as solve a very complex problem. Nonetheless, the possibility remains that a division may be able to devise successful cheating methods without either the information or the calculations involved. Jennergren found that operating rules could be interpreted for individual gain without detection, and consequently without punishment.

Jennergren and Müller [44] performed simulation studies on three kinds of resource allocation schemes. Using an index of cheating to reflect roughly the proportion of the feasible region revealed, they found for the D–W approach that divisions could successfully cheat. For a two divisional setup, if one division cheated, then both divisions had larger payoffs. If both cheat, the result is not so clear—both may obtain higher payoff, or not. In any event, the corporation as a whole is worse off—i.e. corporate profits are lower. However, these losses may not be large. In general, they found that non-cheating is the preferred situation. Minor cheating leads to solutions close to optimal. But heavy cheating (reveal only a small part of production possibilities) will seriously cut corporate returns, but also increase divisional returns markedly. Indeed, this is not a happy situation as the returns to cheating are high. They did not explicitly address the detection issues, but clearly the situation could obtain without detection.

Ljung and Selmer [49] made simulation experiments on real world data using the D–W method in the same way as [44]. They also concluded that it is possible to cheat and the overall profit decreases as cheating increases, and the profit for the division increases at the same time. They observed that small divisions (in terms of profit) gained more by cheating than large divisions. This seems reasonable if we recall the general result that overall profit goes down by cheating. This result has implications for model formulation which we shall discuss in a later section. Briefly, it suggests that divisions in the multilevel firm should be kept relatively large.

Schiff and Lewin [70] in a field study have investigated the cheating issue from a somewhat different view. They were interested in whether divisions would create and husband organizational slack (see [19]), and whether they are likely to be detected. They found divisions can create organizational slack (one situation is the failure to reveal the complete set of production possibilities as discussed above). They are motivated to do so as the usual divisional payoff function is enhanced—and they are not likely to be detected. Thus, there are no

organizational mechanisms driving the system to a more efficient equilibrium.

In terms of multilevel hierarchical models for the firm, the extent of investigation concerning cheating, control, and motivation is sparse. Initial studies suggest it is a serious issue worthy of investigation as the whole scheme is seriously questioned if enforcement can not be obtained for the optimal operating decisions.

Uncertainty and motivation

Uncertainty is a situation where some relevant outcome is determined by chance, e.g. the price for the product, output of a production process, etc.[2] The calculation of optimal decisions for the multilevel situation are somewhat more involved. Secondly, and perhaps more importantly, the distribution of risk bearing is a new issue, which has not been discussed previously. In general, the issue is how to create incentives for the divisional managers in a way consistent with the corporate risk attitude.

The work on the calculation of operating rules under risk neutrality has been investigated by Jennergren [39, 41]. His model involves the allocation of scarce corporate resources to the divisions when it is assumed that allocation must be feasible for each division for all states of nature. A modified D–W decomposed algorithm is devised to solve the problem. Obel and Vander Weide [62] generalized Jennergren's approach using the general version of D–W given in [36]. They showed that even if there were differences in the risk attitudes of the participants, the price–directive method could be used to correctly allocate corporate resources. One property of the model is that top management shares some business risk with the division managers, e.g. if the headquarters is forcing a division to accept a very risky project then the price for the corporate resources used for this project may be lowered. Their specific model is a capital budgeting problem.

The literature on uncertainty and motivation is rather sparse and it is an area for considerable research effort.

Innovation and motivation

Innovation is restricted here to denote the development of more efficient operations, i.e. one unit of a given product can be produced with less quantity of one, or more input resources. Motivation concerns whether it is advantageous to seek innovation and whether it is advantageous to report it when developed. To ask the question if it should reveal the innovations is in theory the same as asking a cheating division if it should stop cheating.

Innovation will increase the profit 'pie', but the multilevel mechanism does not guarantee that the innovating division will obtain the same (or larger) proportion of the pie. Thus, it is possible that divisional profits will decrease under innovation of its own. Briefly, as innovations are introduced, the value

[2] Both the cheating situation and uncertainty are situations of partial ignorance. Cheating is the situation where information is purposefully and deliberately misrepresented. Uncertainty involves the outcome of a random process.

(and cost) of scarce corporate resources will alter, which may lower divisional profits. Ljung and Selmer's [49] study suggests that small divisions will have returns from cheating whilst large divisions will obtain returns from innovation.

V. ORGANIZATIONAL DESIGN AND EMPIRICAL QUESTIONS

The organizational design question is to choose the structure (decomposition) of the firm and the coordination mechanism in order to obtain an optimal solution from a managerial point of view.

The usual definition of optimality is to maximize revenue minus cost ignoring managerial costs. This definition is important from the managerial viewpoint, but not a universal necessity. For example, an organization may not pursue the process to optimality (even if it is possible) and it may be much more important to obtain good improvements in a short time than ultimate optimality. When balancing the marginal cost of managerial efforts with the marginal opportunity cost of lost revenue or increased material costs, it may be a moot point whether the process ever obtains optimality as defined in the mathematical sense. This suggests that the rate of convergence plays an important role in evaluating the structure and the coordinating mechanism.

For most procedures (pricing or budgeting) the numerical experiments have shown that the fewer corporate constraints the higher rate of convergence, and if the number of corporate constraints was fixed the more divisions (departments) the higher rate of convergence. In a managerial context, few corporate constraints means that interactions between the units in the firm will be small, whilst more divisions means that the headquarters should be provided with increasing information. Additional experiments have shown that the starting strategy itself is an important element for the rate of convergence [17].

The question now arises that, for a given firm, is it possible to create a model (and by that the structure of the firm) such that the number of common constraints is minimized and the number of division is maximized subject to a restriction that the structure of the corporate constraints should allow an efficient starting strategy? This problem was treated in [59] where it is shown that the number and the structure of the corporate constraints could be changed by the following instruments:

1. changing the basic structure, whether functional form or the multi-divisional form;[3]
2. combining more divisions into one, so that some corporate constraints become divisional constraints (see [50] for a discussion of the concept 'optimal degree of decomposition');
3. aggregating variables and/or constraints;

[3] See [5] for a discussion on changes in cooperative market relations and their effect upon the basic organization structure.

4. changing the length of the planning periods;
5. removing redundant information;
6. splitting up a corporate constraint into divisional constraints;
7. associating divisions with the headquarters.

Some of these actions may cause suboptimal solutions in the mathematical sense, but the gain from an increase in the rate of convergence may be such that the actions are efficient from a managerial point of view.

As an example, look at the number of divisions—the more divisions the better from an information point of view. But many divisions may cause many corporate constraints. Recall the results about cheating—big divisions did not gain much from cheating, they may even lose. Considering both the computational and managerial reasons, the optimal number of divisions is neither very low nor very high. There exists a tradeoff—more divisions lead to faster computations and more cheating. Thus, the optimal number of divisions could lie in the middle.

A quite different way of changing the model of the firm is to apply a behavioral model (see [24, 64, 65, 66] and the survey in [67]). In these models both the number and the structure of the corporate constraints may differ from a non-behavioral model. These models also introduce more than two-levels in the organizational structure; in the above it has been implicitly assumed that the firm comprises a two-level structure, but in practice this is not likely to be the case. All the methods discussed can be generalized to more than two-levels, but the information flow is no longer a trivial point because different information cycles may exist [74]. It is also possible to use different approaches on different levels; no research has dealt with the best choice of multilevel information cycles, and their implications on practical planning is unknown.

The rate of convergence should however not be the only criterion for choosing the structure and coordinating mechanism. Also relevant may be:

1. the choice of an initial solution to begin the planning process;
2. the rate of adjustment to changing external factors;
3. the speed by which the central and divisional problems can be solved;
4. the accuracy and speed of the communications;
5. the feasibility of intermediate solutions; and
6. the cost of the managerial systems.

Atkins [3] has developed an alternative set of dimensions (or parameters) for comparing multilevel algorithms in organization. The parameters are (p. 620): type of changes, sensitivity to ageing data, guaranteed improvement, justifica-

tion of changes, can consider various alternatives simultaneously and data transmission and location. These parameters are not the same as those above but suggest some concern for equivalent issues. Then he considers both pricing and budgeting approaches with comments on how each algorithm scores on these dimensions.

VI. POTENTIAL FOR PRACTICE

It is clear that there is a gap between the theory of multilevel hierarchical planning systems for the firm and their practical use. However, most large corporations use some method of budgeting for planning and control in a decentralized system. So why is there a gap?

First, the literature has dealt more with computational aspects than with decentralization. This is changing and papers have recently appeared where the algorithmic approaches have been used to discuss organizational design, problems with the location of information, etc. Secondly, the methods have been too naive. They do not model planning in a realistic manner. To illustrate; if we use a resource-directive method then, in the first few iterations (which is the most important in a planning context) the method is very likely to make extreme and unrealistic allocations of the corporate resources by allocating most of the resources (if not all) to the divisions with the highest shadow prices. In the budgeting procedures used in practice the budgets are usually revised only by a small amount. Now the theoretical and practical procedures may be combined by using the algorithm, *but* in the co-ordination, additional restrictions on the changes are put on the reallocations. This combines the reality of practical planning with the power of the algorithmic approaches.

These planning procedures were found to be subject to cheating, poor motivators for innovations, and not well understood in the face of uncertainty. These are important issues—but they need not deter the implementation of the procedures. These ills are also present in the planning systems currently used in practice. Thus, the multilevel planning systems can be used without loss of a benefit obtained in current practice. These methods must be implemented before they can finally be judged. However, if practical aspects are taken into consideration, decentralization by means of algorithmic based planning procedures will be a powerful tool in controlling large organizations. The best next step is to obtain the potential return of optimal seeking algorithmic procedures appended to current planning procedures.

VII. CONCLUSIONS

This paper developed a number of issues concerning the multilevel approach to the organization of the firm. Two main topics were pursued:

1. a review of decomposition of linear models of the firm by pricing and budgeting approaches;
2. a managerially oriented development of important dimensions and questions for empirical investigation concerning multilevel management technology.

A goal for this paper was to review what has been done, but, more important, to suggest that there are important areas for future research in the multilevel models of the firm. The development of empirically valid and operationally meaningful statements concerning trade-offs among different multilevel approach requires further attention. Further, the organization theorist and operations researcher can make the multilevel models relevant to wider audiences for real world application.

The conclusion to this paper is that the multilevel approaches has contributed substantially to the modern theory of the firm along many of its important dimensions. Further, those considerations are not completed and further developments are possible, and *needed.*

l. 5
77

REFERENCES

1. ABADIE J and SAKAROVITCH M (1970) Two methods of decomposition for linear programming. *Proceedings of the Princeton Symposium on Mathematical Programming* (KUHN HW, Ed.). Princeton University Press.
2. ALMON C (1963) Central planning without complete information at the center. *Linear Programming and Extensions* (DANTZIG GB, Ed.). Princeton University Press.
3. ATKINS D (1974) Managerial decentralization and decomposition in mathematical programming. *Opl. Res. Q.* **25**(4), 615–624.
4. BALAS E (1966) An infeasibility-pricing decomposition method for linear programs *Ops Res.* **14**(5), 847–877.
5. BALIGH HH and BURTON RM (1976) Organization structure and cooperative market relations. *Omega* **4**(5), 583–593.
6. BARNARD CI (1938) *The Functions of the Executive.* Harvard University Press. (30th anniversary edition, 1968).
7. BAUMOL WJ and FABIAN (1964) Decomposition, pricing for decentralization and external economies. *Mgmt Sci.* **11**(1), 1–32.
8. BEALE EML, HUDGE PAB and SMALL RE (1965) Experience in using a decomposition program. *Comput. J.* **8**, 13–18.
9. BELL EJ (1965) Primal–dual decomposition programming. *Report ORC* 65-23, Operations Research Center, University of California, Berkeley, California.
10. BEN–ISRAEL A and ROBERS PD (1970) A decomposition method for interval linear programming. *Mgmt Sci.* **16**(5), 374–387.
11. BENDERS JF (1962) Partitioning procedures for solving mixed variable programming problems. *Numerische Mathematik* **4**, 238–252.

12. Bensoussan A (1972) Price decentralization in the case of interrelated payoffs. *Techniques of Optimization* (Balakrishnan AV, Ed.). Academic Press, New York.
13. Burton RM, Damon WW and Loughridge DW (1974) The economics of decomposition: resource allocation vs. transfer pricing. *Decis. Sci.* **5**(3), 297–310.
14. Chadda SS (1967) A decomposition principle for fractional programming. *Opsearch* **4**(3), 123–132.
15. Chandra S (1968) Decomposition principle for linear fractional functional programs. *Revue fr. Informatique Rech. Op.* **2**(10), 65–71.
16. Charnes A, Clower RW and Kortanek KO (1967) Effective control through coherent decentralization with preemptive goals. *Econometrica* **35**(2), 294–320.
17. Christensen JA and Obel B (1976) Simulation of decentralized planning in two Danish organizations using the decomposition scheme from linear programming. Working paper, Institute for History and Social Science, Odense University.
18. Cobb RM and Cord J (1970) Decomposition approaches for solving linked programs. *Proceedings of the Princeton Symposium on Mathematical Programming* (Kuhn HW Ed.). Princeton University Press.
19. Cyert RM and March JG (1963) *A Behavioral Theory of the Firm*. Princeton-Hall.
20. Dantzig G and Wolfe P (1961) The decomposition algorithm for linear programming. *Econometrica* **29**(4), 767–778.
21. Dantzig G and Wolfe P (1960) Decomposition principles for linear programs. *Ops Res.* **8**, 101–111.
22. Downs A (1967) *Inside Bureaucracy*. Little, Brown, Boston.
23. Freeland JR (1975) A note on a resource directive algorithm for allocation of resources in a decentralized organization. *Decis. Sci.* **6**(1), 186–189.
24. Freeland JR and Baker NR (1975) A goal partitioning procedure for modeling coordination activities in a hierarchical decentralized organization. *Omega* **3**(6), 673–688.
25. Freeland JR and Moore JH (1976) Some organizational design implications of resource directive allocation mechanisms. *Research Paper No.* 289, Graduate School of Business, Stanford University (to appear in *Mgmt Sci.*).
26. Geoffrion AM (1970) Primal resource-directive approaches for optimizing nonlinear decomposable systems. *Ops Res.* **18**(3), 375–403.
27. Geoffrion AM (1972) Generalized benders decomposition. *J. Optimization Theory Appl.* **10**(4), 237–260.
28. Glassey CR (1972) Dynamic linear programs for production scheduling. *J. Ops Res. Soc. Am.* **19**(1), 45–56.
29. Glassey CR (1973) Nested decomposition and multi-stage linear programs. *Mgmt Sci.* **20**(3), 282–292.
30. Hass JE (1968) Transfer pricing in a decentralized firm. *Mgmt Sci.* **14**(6), 310–331.
31. Hass JE (1972) Decomposition processes and their use in joint decision making. *Interorganizational Decision Making* (Tuite M, Chrisholm R and Radnor M, Eds.). Aldine, Chicago.
32. Heal GM (1969) Planning without prices. *Rev. Ec. Studies* **36**, 346–362.
33. Heal GM (1971) Planning, prices and increasing returns. *Rev. Ec. Studies* **38**, 281–294.
34. Heal GM (1973) *Theory of Economic Planning*. North Holland, Amsterdam.
35. Ho JK and Manne AS (1974) Nested decomposition for dynamic models. *Math. Prog.* **6**(2), 121–140.
36. Holloway CA (1973) A generalized approach to Dantzig–Wolfe decomposition for concave programs. *Ops Res.* **21**(1), 210–220.
37. Hurwicz L (1972) On informational decentralized systems. *Decision and Organization.* (McGuire CB and Radner Eds.). North Holland, Amsterdam.
38. Hurwicz L (1973) The design of mechanisms for resource allocation. *Am. Ec. Rev.* **LXIII** (2), 1–30.
39. Jennergren LP (1971) Studies in mathematical theory of decentralized resource-allocation. PhD Dissertation, Stanford University.
40. Jennergren LP (1972) Decentralization on the basis of price schedules in linear decomposable resource-allocation problems. *J. Fin. Quant. Analysis* **7**, 1407–1417.

41. JENNERGREN LP (1973) A Note on a Dantzig–Wolfe Decomposition Like Method for Solving a Particular Resource-Allocation Problem under Uncertainty. *Math. Op. forsch. und Statist.* **4**(2), 127–132.
42. JENNERGREN LP (1973) A price schedules decomposition algorithm for linear programming problems. *Econometrica* **41**(5), 965–980.
43. JENNERGREN LP (1975) Decentralization in organizations. *Handbook of Organizational Design* (STARBUCK WH Ed.). Elsevier, Amsterdam.
44. JENNERGREN LP and MÜLLER W (1973) Simulation experiments of resource-allocation decisions in two-level organizations. *Soc. Sci. Res.* **2**(4), 333–352.
45. KORNAI J (1973) Thoughts on multi-level planning systems. *Multi-Level Planning; Case Studies in Mexico* (GOREUX LM and MANNE AS Eds.). North Holland, Amsterdam.
46. KORNAI J and LIPTAK T (1965) Two-level planning. *Econometrica* **33**(1), 141–169.
47. KUTCHER GP (1973) On decomposing price-endogenous models. *Multi-Level Planning; Case Studies in Mexico* (GOREUX LM and MANNE AL Eds.). North Holland, Amsterdam.
48. KYDLAND F (1975) Hierarchical decomposition in linear models. *Mgmt Sci.* **21**(9), 1029–1039.
49. LJUNG B and SELMER J (1975) *Samordnad Planering i Decentraliserade Företag.* PhD Dissertation, Bonniers, Stockholm.
50. MADSEN OBG (1975) A Numerical Comparison of Different Decomposition Algorithms, presented EURO 1 conference, Brussels, Belgium.
51. MALINVAUD E (1967) Decentralized procedures for planning. *Activity Analysis in the Theory of Growth and Planning* (BACHARACH MOL and MALINVAUD E, Eds.). Macmillan, York.
52. MALINVAUD E (1968) Notes sur l'etude des procedures de planification. *Can. J. Ec.* **1**, 16–36.
53. MARGLIN SA (1969) Information in price and command systems of planning. *Public Economics* (MARGOLIS J and GUILLON H Eds.). Macmillan, New York.
54. MARSHAK T (1959) Centralization and decentralization in economic organizations *Econometrica* **27**(3), 399–430.
55. MARSHAK T (1965) Economic theories of organization. *Handbook of Organizations* (MARCH JG Ed.). Rand McNally, Chicago.
56. MESAROVIC MD, MACKO D and TAKAHARA Y (1970) *Theory of Hierarchical Multilevel Systems.* Academic Press, New York.
57. MOESEKE PV and GHELLINCK G (1969) Decentralization in separable programming. *Econometrica* **37**(1), 73–78.
58. MORRIS JR (1975) An application of the decomposition principle to financial decision models. *J. Fin. Quant. Analysis* 37–65.
59. OBEL B (1976) On Organizational Design—from a Linear Programming Point of View. G.S.B.A. working paper no. 183, Duke University.
60. OBEL B A note on a weitzman like planning procedure. *Scand. J. Ec.* (to appear).
61. OBEL B and CHRISTENSEN JA (1976) An Analysis of Procedures for Decentralized Planning, Institute for History and Social Science, Odense University (in Danish).
62. OBEL B and VANDER WEIDE J (1976) A Note on the Decentralized Capital Budgeting Problem Under Uncertainty. G.S.B.A. Working Paper No. 182, Duke University.
63. PEARSON JD (1966) Duality and decomposition technique. *SIAM J. Control* **4**(1), 164–172.
64. RUEFLI TW (1971) Behavioral externalities in decentralized organizations. *Mgmt Sci.* **17**(10), 649–657.
65. RUEFLI TW (1971) A generalized goal decomposition model. *Mgmt Sci.* **17**(8), 353–363.
66. RUEFLI TW (1971) PPBS—an analytical approach, in *Studies in Budgeting* (BYRNE *et al.*), p. 161–180. North-Holland, Amsterdam.
67. RUEFLI TW (1974) Analytical models of resource allocation in hierarchical multi-level systems. *Socio-Ec. Planning Sci.* **8**, 353–363.
68. ROSS M (1969) Two-stage planning in the irish context. *Studies in Regional Science* (SCOTT AJ Ed.). Pion Ltd, London.
69. SANDERS JL (1965) A non-linear decomposition principle. *Ops Res.* **13**, 266–271.
70. SCHIFF M and LEWIN AY (1970) The impact of people on budgets. *Acc. Rev.* XLV (2), 259–268.

71. SCHLEICHER S (1971) Decentralized optimization of linear economic systems with minimum information exchange of the subsystems. *Zeitschrift für Nationalökonomie* **31,** 33–44.
72. SILVERMAN GJ (1972) Primal decomposition of mathematical programs by resource allocation. *Ops Res.* **20,** 58–93.
73. TEN KATE A (1972) Decomposition of linear programs by direct distribution. *Econometrica* **40**(5), 883–898.
74. WEITZMAN M (1970) Iterative multilevel planning with production targets. *Econometrica* **38,** 50–65.
75. WHINSTON A (1964) A dual decomposition algorithm for quadratic programming. *Cah. Cent. Etudes Rech. Op.* **6,** 188–201.
76. WHINSTON A (1964) Price guides in decentralized organizations. *New Perspectives in Organization Research* (COOPER, LEAVITT and SHELLY Eds.). John Wiley, New York.
77. WHINSTON A (1966) A decomposition algorithm for quadratic programming. *Cah. Cent. Etudes Rech. Op.* **8,** 112–131.
78. WILLIAMS HP and REDWOOD AC (1974) A structured linear programming model in the Food Industry. *Opl Res. Q.* **25**(4), 517–527.
79. ZSCHAU EVW (1967) A Primal Decomposition Algorithm for Linear Programming. PhD Dissertation, Stanford University.

ADDRESS FOR CORRESPONDENCE: *Børge Obel, Esq, Institute of History and Social Sciences, Odense University, Niels Bohrs Alle, DK-5000 Odense, Denmark.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

Some Methodological Aspects of Innovation Research

R ROTHWELL

J TOWNSEND

University of Sussex

M TEUBAL

P SPILLER

Falk Institute for Economic Research, Jerusalem

(*Received November* 1976; *in revised form January* 1977)

l. 5
77

From the policy point of view it is essential to understand the links between firm and management characteristics and the outcome of the firm's innovative endeavours, i.e. success or failure. This is achieved via firm behaviour variables, and the causal links between variables at all three levels (characteristics, behaviour, outcome) are demonstrated using interelated sub-systems of variables derived from empirical data obtained from the SAPPHO results.

1. INTRODUCTION

PROJECT SAPPHO was a systematic attempt to discover differences between successful and unsuccessful innovations in the chemical process and scientific instrument industries [1]. The SAPPHO technique involved the detailed comparison of 'paired' successful and unsuccessful innovations in each sector: comparison between pairs was made using a large number of 'project execution' variables, and the method was designed to eliminate the influence of exogenous or 'area' variables, although significant inter-industry differences were explained in terms of structural differences between the two sectors and their respective markets [2]. The Falk research concentrated on the study of successful and unsuccessful innovations in the Israeli biomedical electronic instrument industry: in this case 'project selection' variables figured strongly and area variables were explicitly included in the analysis [3]. Thus, while SAPPHO was basically a study of the *management of innovation*, Falk was an attempt at identifying *comparative advantage* in innovation.

This paper describes how, using a number of concepts developed by the Falk researchers, the very detailed SAPPHO data were re-analysed in such a way that area variables could be accommodated and inter-industry differences better explained. A variable typology was developed and, based on this typology, a number of theoretical 'sub-systems' of interrelated variables were postulated: an empirical test was then made which provided some validation of these sub-systems.

2. AREA VARIABLES

It is, of course, unrealistic to pretend to understand the dynamics of innovation in a particular industry sector without understanding something of the characteristics of that sector and of the markets it operates in. In order to facilitate the better understanding of the relationship between the characteristics of a new product or process, the innovating organisation and the market place, a number of area variables have been defined, namely 'market determinateness' and 'technological determinateness'. The first of these was used by the Falk researchers in interpreting comparative advantage between firms in the Israeli biomedical instrument sector [3].

Market determinateness

This variable indicates the extent to which user needs have been translated into product classes, functions and specifications. For example, the need to

TABLE 1. SPECIFICITY OF SIGNALS AND MARKET DETERMINATENESS

Degree of market determinateness	Type of signal (√) or type of demand anticipation (×)			
	About need (1)	About product class (2)	About basic functions* (3)	About product specification* (4)
1. No clearly perceived general need by user	×	×	×	×
2. Clearly perceived need, but product unknown	√	×	×	×
3. Product known, but basic function not completely standardised	√	√	×	×
4. Product known and basic function standardised	√	√	√	×
5. Commissioned R&D programme	√	√	√	√

* In the context of biomedical electronics each one of the physiological parameters measured in, for example, a cardiovascular monitoring system (ECG, blood pressure, etc.) would represent a *basic function*. *Product specifications*, on the other hand, would determine the efficiency, reliability, simplicity of operation, and degree of patient safety with which the basic functions are performed.

reduce costs by 20% is a very general need with no expression in terms of products or specifications. When the user commissions a particular type of cost-reducing device with certain characteristics, however, the need is very specific. In the first case there is a high risk of failing to understand the precise structure of user needs, which in any case may not be very well defined. Moreover, the product-acceptance risk due to the user's incomplete information on the performance characteristics of the product is likely to be very high. These risks diminish when market determinateness increases.

A formulation of this variable in terms of signals, and a translation of the theoretical variable into categories appropriate to biomedical instrumentation, are given in Table 1. This shows clearly that market determinateness is a measure of the degree of specificity of the signals deriving from the market place.

Technological determinateness[1]

The concept of market determinateness suggests a parallel area variable pertaining to technology, one indicating the extent to which existing science and technology determine specific solutions to the technical problems posed by the innovation. Within a particular area, technological determinateness is inversely related to technical risk, while technical risk, together with R&D effort, firm's familiarity with technology, etc. would determine efficiency in development and lead time.

3. TYPOLOGY OF VARIABLES

The variables used in Project SAPPHO suggest a hierarchy of at least three levels: I. manifestations or results of firm behaviour; II. variables related to firm behaviour; III. firm and management characteristics. SAPPHO stated, for example, as one of the prime differences between successful and unsuccessful innovators, 'successes understood user needs better than did failures'. Failure to understand user needs might be a result of a lack of search on user needs, which in turn, might be a consequence of a technical bias on the part of management. Thus 'failure to understand user needs', which is a *manifestation of firm behaviour*, can be linked to the *management characteristic* 'technical bias of management' via the *firm behaviour* variable 'lack of search on user needs'.

The distinction between firm behaviour and firm and management characteristics is important from the point of view of policy since, while firm behaviour can, in principle, change in the short term, firm and management characteristics may only be amenable to gradual change. It is also of paramount importance to identify those characteristics of firms and management which are relevant for

[1] This concept has yet to be developed fully.

success in innovation, which requires an explanation of the link between these characteristics and the results of firm behaviour: any such explanation should be expressed in terms of firm behaviour variables.

A three-level hierarchy of variables such as that suggested above should prove a valuable aid towards a better understanding of, and explanation of, success and failure in innovation. A classification of some of the most significant SAPPHO and Falk variables is presented in Table 2.

TABLE 2. VARIABLE TYPOLOGY

Level I Manifestations or results of firm behaviour
Understanding user needs
Required adaption by users of the innovation
Efficiency in development
After-sales problems
Lead time
First to market
Level II Variables related to firm behaviour
a. *Programme execution*
Search on user needs
Customer involvement in development
Promotion and marketing effort
User education
Planning of R&D
R&D effort
Spectrum and/or line offered
Innovation profile
External coupling in specific technological areas
b. *Programme selection*
R&D synergy with other programmes
Production synergy with other programmes
Marketing synergy with other programmes
Integration with overall firm strategy
Area variables
Market determinateness
Technological determinateness
Radicality for world technology
Strength of competition
Type of competition
Rate of growth of market
c. *Overall firm behaviour*
Firm's portfolio of innovations
Share of R&D in firm's budget
Level III Firm and management characteristics
Firm size
Characteristics of firm organisation (flexibility)
Internal and external communication
Management orientation
Characteristics of programme managers (authority, commitment, and diversity of experience)
R&D strength of firm

Vol. 5
1977

4. POSTULATING SUBSYSTEMS OF VARIABLES

The above typology suggested a set of links among variables at different levels. This has implications for the empirical test of the significance of particular variables or measures for performance. Level I variables should be directly correlated with preformance, while each variable of the sets of lower level variables should be correlated with those at higher levels. (This contrasts with both Falk and SAPPHO in that only *direct* association between each variable and performance were tested.) It was thought possible to test the significance of variables belonging to more than one classification by postulating a number of *variable sub-systems* of possible relevance for performance and try to test the significance of each.[2] This would be a structural approach and would need a theoretical framework.

The results of the SAPPHO and Falk projects indicate the potential relevance of two sub-systems of linked variables: *A* which might be called the *user-needs*

TABLE 3. TWO SUB-SYSTEMS

Sub-system *A*	Sub-system *B*
Level I	*Level* I
Understanding user needs	First to market (or lead time)
Level II	*Level* II
Search on user needs	R&D effort
User education*	Investment in first fullscale plant
Marketing*	R&D synergy†
Firm's market involvement*†	Production synergy†
Level III‡	*Level* III‡
External communication	R&D strength

* These variables are especially relevant in a broad view of the user need sub-system where both the determination of user needs *and* the generation of product acceptance are crucial. A good example is bio-medical electronics instrumentation [3].

† These variables result from the intersection of innovation and firm characteristics and can be traced as resulting from firm behaviour (programme selection, as distinct from programme execution).

‡We have not explicitly written firm size and management characteristics separately for each sub-system because they are relevant variables for both. The specific characteristics of management may, however, vary from one sub-system to another, the determination of which ones may be the object of specific tests.

[2] A variable sub-system comprises a subset of variables and their pattern of inter-connection selected on the assumption that it dominates other sub-systems in its potential for explaining performance.

sub-system and *B* which might be called the *lead-time* or *first-to-market* sub-system. The former was seen to be particularly relevant for SAPPHO's and Falk's instrument innovations;[3] the latter was relevant for SAPPHO's chemical innovations.[4] Other sub-systems might also be suggested, but most seem to be less relevant. Sub-systems A and B, expressed in terms of the variables and levels presented in Table 2, are shown in Table 3.

The relevance of each sub-system in an analysis of performance depends on the general conditions surrounding the universe of innovations being tested. These general conditions are reflected by area variables. Our presumption is that the lead-time sub-system is the relevant one if the various competing innovations are reasonably well adapted to user needs. Only then can first-to-market (and lower cost) be critical factors for success. The conditions for this to occur are strong market determinateness and strong technological determinateness. The former guarantees that the understanding of user needs presents few problems, while the latter ensures that there is little technical problem in translating user needs into the product.

Whenever market or technological determinateness (or both) are weak, there is reason to believe that the competing products will differ in their degree of adaptation to user needs. Our presumption is that if market determinateness is weak and technological determinateness strong, the user-needs sub-system will be dominant; here there will be an unequal perception of user needs by innovating firms.[5]

Vol. 5
1977

Most of the SAPPHO chemical innovations were process innovations not including fundamental changes in existing products, while most instrument innovations involved a significant new product component. It is relatively easy to formulate a precise set of specifications for a new chemical, and once the specifications of a chemical substance are established, there are usually only a few alternative reactions which may be developed to achieve these specifications. Therefore, R&D effort expended and lead times would be important factors in performance. Instrument innovations, on the other hand, imply a wider feasible set of alternative product specifications (including size, functions, versatility) relative to those technologically feasible in the obtention of new chemical substances. The possibilities of coupling technology with needs are, consequently, greater in instruments, and the extent by which this is done will vary more widely among the instrument firms than among chemical firms.

[3] The SAPPHO results for the instrument industry showed a very much stronger relationship between market factors (search on user needs, user education, marketing effort, firm's involvement in the market) and success and failure, than did the results for the chemical industry.

[4] In the chemical industry, most successes were first-to-market. In instruments, and opposite pattern was found.

[5] In the opposite case (i.e. high market determinateness and low technological determinateness), the high technical risk will very likely lead to varying degrees of translating clearly perceived user needs into a product. In this case an efficiency-in-development sub-system would be dominant.

The above suggests that market determinateness is higher for chemical innovations than for instrument innovations. This may explain the smaller role of R&D strength and lead times in performance in instruments. In fact, lead time may only be reasonably well defined when market determinateness is high. The innovations in chemicals also appear to be much more technologically determinate than the instrument innovations. For these reasons we expect the user-needs sub-system to be most important in the instrument sector and the lead time sub-system to be dominant in chemicals.

5. A PILOT TEST

A pilot test of the sub-system approach was conducted for 12 SAPPHO pairs from each sector. The specific object was to test the relative strength of each of the two sub-systems in each sector. The set of variables chosen was determined by the need to test the relevance of each of the sub-systems postulated. A matrix of interconnections was used to determine the relevance of particular sub-systems for performance, which summarizes the associations between the various variables across the innovation set. Since the theoretical notion of hierarchies of variables provided a structure to the kind of interconnections sought, only those between variables at adjacent levels were examined.

Two matrices were built, one for each industry. As an example of how the links were correlated, if the successful innovator of a pair 'understood user needs better' *and* 'had more user involvement' than the failed member, the pair

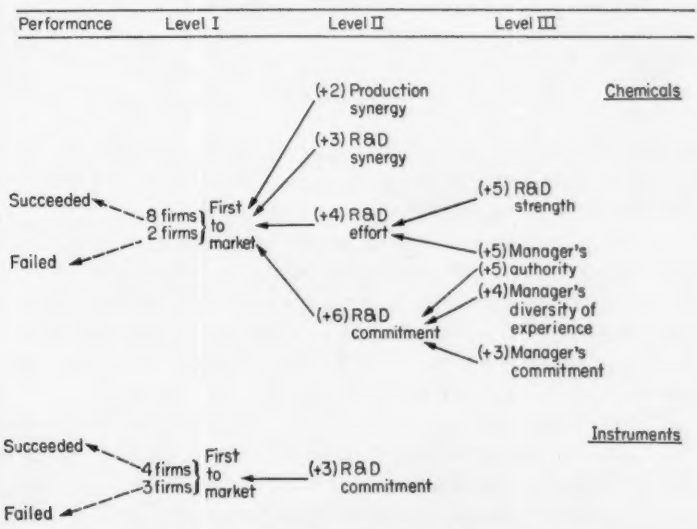


FIG. 1. *Lead-time sub-systems; both industries.*

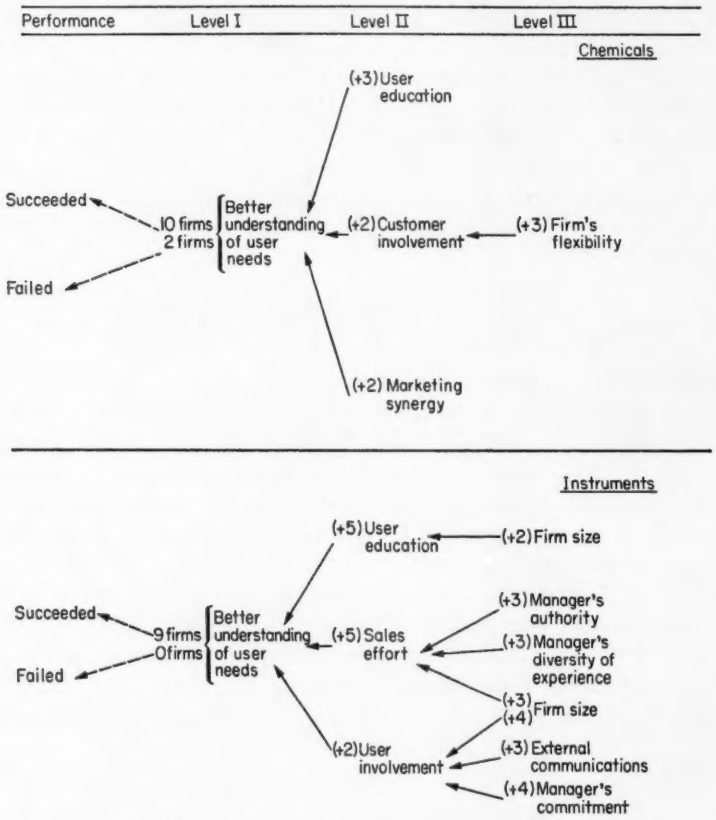


FIG. 2. *User-needs sub-system; both industries.*

was said to show a positive correlation between these two variables (recorded as +1): if the ordinal relation was in the opposite direction, −1 was entered: no correlation (0) included ambiguous cases and lack of information.

Since the extraction of sub-systems from the matrices was governed by theoretical considerations, this meant that impossible or illogical interactions, however strong, were rejected. For example, the matrix for the chemical industry showed a strong relationship between 'sales effort' and 'R&D strength'. Now, both these might be linked via, for example, firm size (more resources available), but any direct causal relationship between the two can be discounted.

Figures 1 and 2 show the lead time and user-needs sub-systems respectively for both industries which were derived from the matrices. It can be seen that, as predicted, the lead-time sub-system is dominant in chemicals (and has little structure in instruments) while the user-needs sub-system is dominant in instruments (and has little structure in chemicals). Moreover, both the dominant sub-

Vol. 5
1977

systems are highly structured, and both show significant relationships between Level III and Level II, and Level II and Level I variables. (The numbers show the strength of the links.)

6. DISCUSSION

We have demonstrated (albeit with a limited test) using the sub-system concept, linked with a three tier variable typology, and in the context of a suitable set of area variables, that it is feasible, from empirical data, to identify sub-sets of linked variables at different levels, which increased our ability to explain success and failure in innovation. For example, while SAPPHO offered 'a better understanding of user needs' as the prime factor differentiating for success, the user-needs sub-system for the instrument pairs (Fig. 2) showed the links between achieving this superior understanding of user needs, and firm and management characteristics (e.g. firm size, manager's authority); these links were the firm behaviour characteristics (e.g. user education, sales effort). Further, the area variables market determinateness and technological determinateness provided an explanation for the (predicted) dominance of the user-needs sub-system in instruments, and the lead time sub-system in chemicals.

One of the more important aspects of the sub-system approach is that it identified higher level factors which, while not *directly* correlated with performance, nevertheless played a significant (indirect) role in success or failure. The results for the SAPPHO instrument industry, for example, showed no correlation between firm size and success or failure: the instrument user-needs sub-system showed, however, (Fig. 2) a relationship between firm size and the three level II variables, 'user education', 'sales effort' and 'user involvement'; these were, in turn, related to the level I variable 'a better understanding of user needs' which was, itself, very strongly correlated with success. This firm size–success relationship was not, as stated above, evident from the original SAPPHO analysis; the user-needs sub-system, however, clearly demonstrated how the links between the two operated.

Most empirical studies of the innovation process offer a list of individual, and often apparently unrelated, factors which are found to be correlated with success or failure. It is, however, unrealistic to expect managers to think of their projects in this manner, and results couched in such terms are of only limited utility to managers. The sub-system approach, which is based on a structured hierarchical classification of variables, and which identifies interactive groups of variables at different levels, does, we believe, offer both researchers and managers a more satisfactory conceptual framework for analysing and understanding the innovation process.

Finally, while the test undertaken in section 5 was only a pilot test performed using a limited data base, nevertheless the results did, we feel, demonstrate the

potential usefulness of the sub-system approach as a powerful analytical tool. We hope that other researchers (and managers) will further test and extend this technique through operating on their own empirical data.[6]

ACKNOWLEDGEMENTS

The SPRU authors wish to acknowledge the financial support of the Science Research Council. The Falk authors wish to acknowledge the financial support of the National Council for R&D, Prime Minister's Office, Jerusalem, and the kind and valuable help from Shosona Freund and Manuel Trachtenberg.

REFERENCES

1. SPRU (1972) *Success and Failure in Industrial Innovation: Report on Project SAPPHO*. CSII, London.
2. ROTHWELL R, FLEEMAN C, HORSLEY A, JERVIS P, ROBERTSON AB and TOWNSEND J, (1974) SAPPHO updated: project SAPPHO phase II. *Res. Pol.* **3**, 258–291.
3. TEUBAL M, ARNON N & TRACHTENBERG M (1976) Performance in innovation in the Israeli electronics industry: a case study of biomedical electronics instrumentation. *Res. Pol.* **5**(4), 354–379.

ADDRESS FOR CORRESPONDENCE: *Dr Roy Rothwell, The University of Sussex, Science Policy Research Unit, Mantell Building, Falmer, Brighton, Sussex BN1 9RF, UK.*

[6] For example, the area variable 'technological determinateness' requires further refinement. Other area variables such as 'strength of competition' (while is especially crucial for an analysis of comparative advantage in innovation) requires careful development, as do 'type of competition', 'rate of growth of market', etc. Other potential sub-systems such as 'efficiency-in-development' requires definition and testing.

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

Contingency Analysis for Effective Administration

PAT JOYNT

Oslo Institute of Business

Received November 1976; *in revised form February* 1977)

Contingency theory has had a strong influence in recent administration analysis and organization theory. Much of the previous work has been limited by the number of contingencies analyzed. This article reports on the findings from 56 Norwegian companies and attempts to analyze the contingencies of effectiveness, efficiency, innovation and satisfaction using the independent variables of environment, structure, technology, communication, motivation, decision making and job design.

CONTINGENCY ANALYSIS FOR EFFECTIVE ADMINISTRATION

SINCE THE 1960's, a major shift in the directions and emphasis in organization and administrative analysis has taken place. This new emphasis is called 'contingency theory', which argues that there is no one best way to organize for effective results, and that the emphasis on results may vary from organization to organization. The early bureaucratic approach emphasized by Taylor, Weber and Fayol (classical, theory *X*, authoritative or mechanistic) was replaced by the human relations approach of the 30's, 40's and 50's (social, theory *Y*, democratic, or humanistic) and this in turn was supplemented with the change or innovation approach of the 60's (modern, group work, shared decision making, or organic). Much of this can be documented in the works and research of Taylor [21], Roethlesberger [20], Weber [26], McGregor [14], Etzioni [3], Burns and Stalker [2], Woodward [27], Thompson [22], Lawrence and Lorsch [11] and more recently Kast and Rosenzweig [9].

Contingency theory essentially tells us that for different organization situations or contingencies, different factors are needed to make for successful administration. A review of the theory and recent research would indicate that the following factors are important determinants in describing and understanding organization from a contingency point of view:

Environment

The environment includes such factors as the stability of the market, the amount of competition and the amount of product changes necessary. Environment was the essential theme of the Lawrence and Lorsch research in *Organization and Environment* [11]. They found that different structures were required for an organization to be successful in an unstable and changing environment from those required in a stable and unchanging environment. The earlier work of Burns and Stalker [2] also support this hypothesis.

Technology

While many definitions of technology exist, one of the most used is Thompson's typology of intensive, service, and long linked technology [22]. An example of intensive technology is a hospital operating room where coordination of a variety of disciplines is necessary in order to ensure effectiveness. A bank might be used as an example of a service institution or service technology, and a car assembly line could be used to illustrate long linked or assembly line technology. Again research shows us that different types of organization structure are required for different types of technology.

Vol. 5
1977

Structure

Structure includes factors such as amount of formal rules, levels in the hierarchy, spans of management, use of goals and size of organization.

Communication

Communication (or types of communication) is often considered as an extension of the structure factor. In addition the use or non-use of computer technology is a very important aspect of this factor. The use of the telephone, letters, notes or other channels along with the direction of communication, be it vertical or horizontal, or internal or external are also important considerations.

Motivation

Motivation is probably the most emphasized variable in the average business school. A review of the past and recent theory and research shows it is extremely difficult to find agreement on the factors and/or the processes that actually influence behavior. The early theories of the content of factors presented by McGregor, Maslow and Herzberg are being replaced by the new process theories of Vroom, Lawler, Skinner, and others [7, 8, 24].

Leadership

The leadership variable has had a rather rough road to follow using the map

of administration analysis. The trait approach, the situation approach, the style approach and the path goal approach have all been well documented in the literature. The problem that always emerges is lack of results, specifically research evidence that would allow one to conclude that one approach is superior to the others. Much of the recent work has involved a path–goal analysis [6] where the leader assists the subordinate along the correct path towards goal achievement by providing: the necessary

support;
interaction help;
clarification of goals;
help in the work situation [1].

Decision making

The use of groups in the decision process has been one of the major themes in the recent behavioral literature. In some situations the single person proves most effective while in others the group may be the best strategy [25].

Job design

Building variety, task identity, autonomy, feedback [10] and a factor that Thorsrud [23] calls 'the need to see the connection between what one does and produces and the firm's goal and the use it has in the market or society as a whole', are all part of a growing agreement on what a job should include if the organization participant is to be satisfied with his job.

In addition to the above determinants, the major emphasis or goals of an organization are required to complete the analytical picture. Much of the early literature concentrated on profit or effectiveness as the primary goal. While this has been the subject of much debate in recent years, behavior in most organizations still has a high orientation towards effectiveness. Etzioni [3] and microeconomic theory have refined this concept to include not only completing a job (effectiveness) but also doing the job using resources as effectively as possible (efficiency). In addition to effectiveness and efficiency the human relations of the 40's and 50's and the industrial democracy movements of recent years have caused many organizations to emphasize morale or more specifically job satisfaction as an end in itself. In Sweden, for example, the recent Volvo and Saab experiments emphasized job satisfaction first and effectiveness as a secondary outcome.

Burns and Stalker's excellent work *The Management of Innovation* [2] introduced the fourth goal in our modern model of contingency theory. With contracting product life cycles and changing consumer needs, many organizations have found it necessary to introduce innovation as a goal in order to survive. The well known Harvard Business Review article 'Beyond Theory *Y*' illustrated the importance of using innovation in a contingency model [15].

CONTINGENCY THEORY FOR THE 70's

Few analyses have attempted to integrate all of the above into a framework for analysis. Possibly the main reason is the magnitude of samples required to make the research meaningful. Another reason might be a preference of researches to emphasise the new concepts of innovation and job satisfaction at the expense of the older concepts of effectiveness and efficiency. In an attempt to correct this discrepancy the following model (Fig. 1) was used in an analysis of administration in 56 Norwegian organizations.

Determinants or independent variables	Organization goals (emphasis) or dependent variables
Environment Technology Structure Communication Motivation Leadership Decision making Job analysis	Job satisfaction Innovation Effectiveness (goal completion) Efficiency (resource use/time)

FIG. 1. *Contingency model for administrative analysis.*

Interviews, historical data and questionnaires were used in the analysis of a variety of organizations which included hospitals, churches, oil companies, newspapers, assembly plants and research centers. The analysis reported in this article is taken from the questionnaire data. An average of 20 participants mainly in management positions from each of the 56 organizations completed a questionnaire which took about an hour to complete. From this data 4 contingency models were generated, each describing one of the four dependent variables mentioned earlier. The results for the most part support many of the individual theories of satisfaction, innovation, effectiveness and efficiency. More importantly they show the necessity of using contingency analysis in explaining modern organization behavior as many unique models emerged in the analysis.

The following figures are based on statistical analysis of the aggregate data which eventually resulted in a regression model for each of the dependent variables. The regression equation is shown at the bottom of each figure.

The following figures are drawn to illustrate the amount of emphasis each factor has. For example, a large amount of time in groups is much more important than the use of intensive technology and thus the larger the box in the figure.

The results presented in Figs. 2–5 present a framework that managers might use to evaluate the goals of their own organization. As an example, if one is

Vol. 5
1977

FIG. 2. *Innovation as an emphasis.*

Regression model:
0·7 *group work* + 0·45 *job feedback* + 0·36 *participation* + 0·3 *dynamic environment* + 0·23 *communication horizontal* + 0·16 *technology intensive* = *organization innovation emphasis.*

Variance explained: 52·5%.

interested in improving efficiency (that is using resources more efficiently) our data suggests that one:

(1) Use informal notes. Notes provide a record as well as being easier to use than the formal letter, telephone or personal contact.

(2) Decision making should be done individually after getting the necessary inputs from subordinates, peers and overordinates.

(3) Job autonomy should be reduced. This suggests the use of job descriptions and formal procedures if efficiencies are desired.

(4) Leadership or management should be socially oriented.

One of the most interesting results was the use of notetaking. Many organizations may be losing efficiency when emphasising that telephones be used, or letters be typewritten, or communication be more personal. But it is also possible one could over-emphasize notetaking at the expense of reducing other goals such as innovation or job satisfaction.

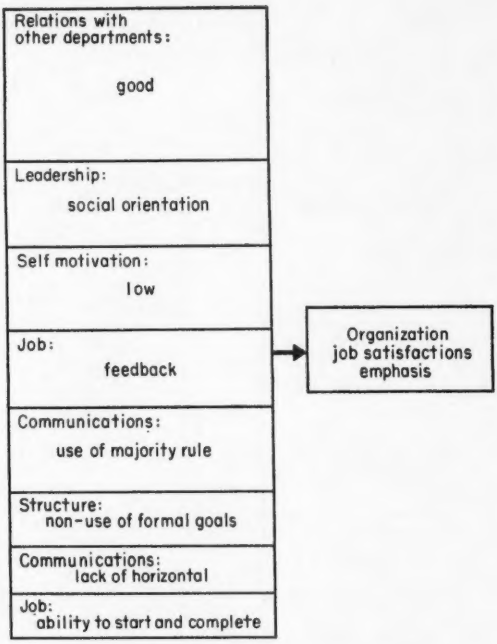


FIG. 3. *Job satisfaction as an emphasis.*

Regression model:
0·60 *relations good with other departments* + 0·39 *leadership social oriented* − 0·38 *motivation high* + 0·3 *job feedback* + 0·24 *communications use of majority rule* − 0·12 *structure use of formal goals* − 0·1 *communications horizontal* + 0·09 *job ability to start and complete task* = *job satisfaction.*

Variance explained: 79·9%.

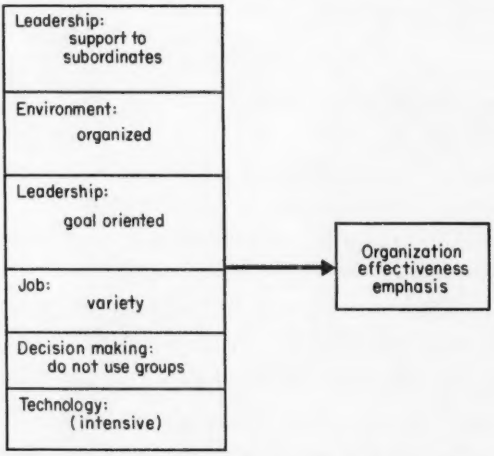


FIG. 4. *Effectiveness (goal or job competition) as an emphasis.*

Regression model:
0·27 *leadership support* + 0·23 *environment organized* + 0·21 *leadership goal oriented* + 0·16 *job with variety* − 0·11 *decision making using groups* + 0·1 *technology intensive* = *organization effectiveness emphasis.*

Variance explained: 56%

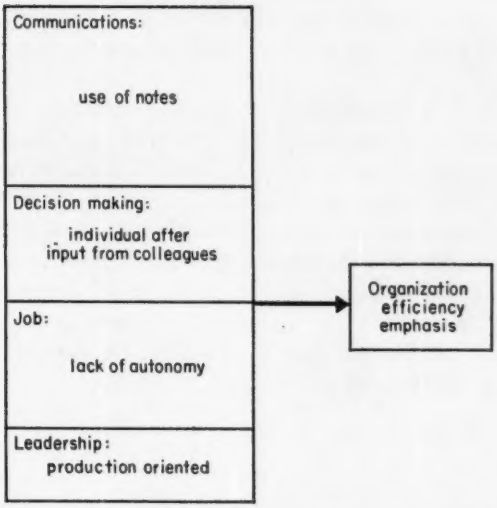


FIG. 5. *Efficiency as an emphasis.*

Regression model:
0·42 *Communications use of notes* + 0·28 *decision making individual after input from colleagues* − 0·25 *job autonomy* + 12 *leadership production oriented* = *organization efficiency emphasis.*

Variance explained: 52·2%.

The most interesting model seems to be the one relating to job satisfaction. This model shows that relations should be good with other departments, that the managers involved should have a social or cooperation orientation, that job feedback is necessary, that majority rule be used, that formal goals be de-emphasized, that a minimum of horizontal communications exist and that each individual should have the opportunity to start and complete a job. All of this agrees with much of the theory on job satisfaction which has been the major emphasis in recent years, particularly in Scandinavia. The autonomous work group would fit extremely well into the above model. But one variable is lacking from existing theories, and that is negative self motivation. Our results show a strong relationship between the individual with a low amount of self motivation and an organizational emphasis on morale or job satisfaction.

There are several possible explanations for this. An emphasis on satisfaction does not occur until after management realizes that self motivation is low, as indicated perhaps by excessive turnover, absenteeism or poor quantity and/or quality in work. An example of this would be the Volvo experiments with autonomous work groups. Two of the major reasons for management seeking a change (or emphasizing satisfaction) were high turnover and absenteeism. Another explanation might be the use of improper techniques in an effort to increase job satisfaction. Many of the earlier strategies for job satisfaction have

produced very mixed results. It is possible that many organizations are using strategies in this respect which result in frustation and lower rather than higher self motivation.

Another explanation might be the research instrument used to obtain the data. Respondents were asked to 'evaluate your own motivation using a 5 point scale from a very high to very low'. This produced a cluster of results at the high end of the scale, suggesting a possible distortion in our results. One also has to ask whether one can evaluate one's own motivation on a five point scale. A fourth explanation may be connected to the types of needs and rewards valued by Norwegian middle managers. Earlier evidence [7] showed that managers have a tendency to evaluate their chances for promotion as poor if they performed well on the job. Further analysis of our present data showed rewards or needs in the following order of desirability.

1. Positive peer relationship;
2. Self respect;
3. Pay;
4. Using one's skills on the job;
5. Professional growth;
6. Self realization;
7. Supervisor relationship;
8. Security;
9. Responsibility;
10. Recognition;
11. Promotion;
12. Social activities;
13. Power;
14. Better work procedures;
15. Prestige.

Group needs such as positive peer relationships and self respect were much stronger than the individual promotion and power motives often associated with high motivation [13]. This suggests the autonomous work group as a viable means of increasing satisfaction rather than the traditional model of using rewards such as promotion, power, and prestige.

The model for innovation agrees very well with the theory and prior research in this area. A dynamic environment, the use of groups, horizontal communications and intensive technologies are well documented. Job feedback has not been specifically mentioned but in a changing organization the need for feedback is essential if change is to progress satisfactorily.

The effectiveness model suggests a routine or organized environment, leadership that provides psychological support as well as having a goal orientation, the limited use of groups in decision making, jobs with variety and technologies of the intensive type where many of the disciplines or tasks require coordination.

These last two factors disagree with much of the prior theory. Generally one would expect to find routinized jobs or assembly line jobs if effectiveness is to be emphasized. Again it is possible that a cultural bias has entered the picture. The emphasis in recent years in Norway has been on increasing variety in jobs. The results in this research would suggest that this strategy has resulted in possible gains in effectiveness for the organizations involved. This has also been the theme in many of the participation experiments in Scandinavia particularly at Volvo where one of the goals of management was to maintain effectiveness with the introduction of the autonomous work groups.

Intensive technologies as part of the model may be another factor unique to Norway. The average company or organization in the sample used is small compared to samples used in similar research endeavours. The emphasis is often on service and intensive technologies and not on assembly technologies. There are no automobile factories and few large assembly line factories in the country and thus a minimum of data were available from these types of organizations. The variable technology was also weak in the model as the regression equation shows.

The above results re-emphasize the need for a complex analysis in understanding organizations. It is also important to note that no organization can be analyzed using only *one* of the above end variables. Every organization emphasizes each of the variables to a different degree. Several examples may serve to clarify this:

A research unit may have a high emphasis on innovation followed by less of an emphasis on satisfaction and even less of an emphasis on effectiveness and efficiency. One could rank order the emphasis of this organization as follows:

1. Innovation;
2. Satisfaction;
3. Effectiveness;
4. Efficiency.

A large assembly plant may have an emphasis of production first, satisfaction next and possibly a negative emphasis on innovation. One might rank the order of the goals of this organization as follows:

1. Effectiveness;
2. Efficiency;
3. Satisfaction.

Note that we have eliminated an emphasis on innovation in the above. One might also find the above to be the situation for one assembly plant while another plant with almost the same characteristics may reverse the order:

1. Satisfaction;
2. Effectiveness;
3. Efficiency.

The above examples complicate some of the prior work in this area in that much of the work emphasized contingency models with only *one* dependent factor or goal. The research unit was classified as innovative, the assembly plant as effective and a church organization might have satisfaction as its main emphasis. Our research shows that in most cases all of the dependent factors, effectiveness, efficiency, innovation and satisfaction are emphasized but in varying degrees. It all depends on:

(a) the technologies involved;
(b) the environment;
(c) the methods of communication, motivation;
(d) leadership and group work;
(e) the types of jobs.

CONCLUSIONS

The results presented strongly suggest that the contingency approach to the study and practice of management is the most appropriate one. "Like other (recent) studies from the contingency perspective, the central message of our findings for managers is that there is no one best solution to the administrative and organizational issues they face. Rather, questions of managerial behavior and organizational design must be considered in the light of the total environment which includes our independent factors previously mentioned. This position is at variance with the broad panaceas about management practices which have been advocated by many early authors. For example, it is questionable whether the early prescriptions of Likert and McGregor with their emphasis on participative leadership and lower level involvement in decisions, would lead to positive outcomes in all situations [12]."

The results also raise another very important question. The traditional textbook emphasis is often on innovation as the major dependent variable along with satisfaction as a close second. The results reported showed effectiveness was the main emphasis for the sample of 56 organizations considered. Perhaps this is a trait peculiar to Norway, but perhaps not. In a sense the results may reopen the door to the human relations, classical school and management of innovation debate. As an example the book *Organization and Environment* by Lawrence and Lorsch [11] as well as the *Management of Innovation* by Burns and Stalker [2] place more of an emphasis on innovation than on productivity. Is it possible that today's management students are getting the wrong preparation for their future roles as managers with texts implying that innovation and satisfaction are the main emphasis?

The implications of contingency theory can be both a comfort and a source of frustration. It is a comfort to those who may have tried the traditional solutions and found they did not work. One can now blame the theory rather

Vol. 5
1977

than the managers or other factors involved. It is a frustration to those who still seek the one right way because contingency theory in its elementary form keeps telling us—"It all depends on".

If so, perhaps the best business school strategy is to educate managers in describing their own environment, i.e. to make researchers of them. Perhaps the essence of management consists in asking the right question and putting the answers in the proper perspective.

REFERENCES

1. Bowers & Seashore (1969) Michigan Center for Research, Ann Harbor.
2. Burns T & Stalker GM (1961) *The Management of Innovation*. Tavistock, London.
3. Etzioni A (1961) *Modern Organizations*. Prentice Hall, Englewood Cliffs.
4. Gustavsen B (1976) *Bedriftsorganisasjon—Allernative Modeller*. Tanum, Oslo.
5. Hage J & Aiken M (1970) *Social Change in Complex Organizations*. Random House, New York.
6. House RJ (1971) A path–goal theory of leadership effectiveness. *Admin. Sci. Q.* **16**, 321–325.
7. Joynt P (1975) *Expectancy Motivation Models*, BI skrift- serie hefte nr. 4, p. 7, Oslo.
8. Joynt P (1976) *Management Concepts*. BI Press, Oslo.
9. Kast F & Rosenzweig J (1970) *Organization and Management: A Systems Approach*. McGraw–Hill, London.
10. Lawler E (1974) For a more Effective Organization—Match the Job to the man. *Organizational Dynamics*, Summer edition, pp. 21–27.
11. Lawrence P & Lorsch J (1967) *Organization and Environment*. Harvard Business School, Division of Research, Boston.
12. Lorsch J & Morse J (1974) *Organizations and their members: A Contingency Approach*. Harper & Row, New York.
13. McClelland DC (1961) *The Achieving Society*. Van Nostrand, Princeton, NJ.
14. McGregor D (1960) *The Human Side of Enterprise*. McGraw–Hill, New York.
15. Morse JJ & Lorsch JW Beyond theory *Y. Harvard Bus. Rev.* **48**(3), 61–68.
16. Norman R (1971) Organizational Innovations, Product Variation and Reorientation. *Admin. Sci. Q.* **16**, 203.
17. Paulson SK (1974) Causal Analysis of Interorganizational Relations: An Axiomatic Theory Revised. *Admin. Sci. Q.* **19.**
18. Porter LW & Lawler EE (1968) *Managerial Attitude and Performance*. Irwin, Homewood.
19. Pugh D, Hickson J, Hennings C & Turner A (1969) The Context of Organizational Structures. *Admin. Sci. Q.* **14.**
20. Roethliesberger F & Dickson W (1939) *Management and the Worker*. Harvard University Press.
21. Taylor F (1911) *The Principles of Scientific Management*. Harper & Row, New York.
22. Thompson JD (1967) *Organziations in Action*. McGraw–Hill, New York.
23. Thorsrud E & Emery F (1964) *Industielt Demokrati: Representasjon på styreplan i bedriftene. Noen Norske og utenlandske erfaringer*. Universitetsforlaget, Oslo.
24. Vroom VA (1964) *Work and Motivation*. Wiley, New York.
25. Vroom VA & Yetton PW (1973) *Leadership and Decision Making*. University of Pittsburg Press, Pittsburg.
26. Weber M (1947) *The Theory of Social and Economic Organization*. The Free Press, Glencoe.
27. Woodward J (1965) *Industrial Organizations: Theory and Practice*. Oxford University Press, London.

ADDRESS FOR CORRESPONDENCE: *Pat Joynt or Øivind Revang, Research Unit, Oslo Institute of Business, Frysjaveien 33 c, Oslo, Norway.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

Management Perspectives in Physical Distribution

SAMUEL EILON

Imperial College, London

(*Received March 1977*)

Distribution problems may conveniently be divided into three categories—strategic, tactical and operational—and although each is bound to be affected by the others, they involve different time scales, different constraints and different requirements for information inputs. At a strategic level, models can be constructed for the rate of return on investment and for unit cost to help identify the major elements that affect performance, and in this way any specific change in given components in the distribution system can be traced and its possible effects can then be identified. Another strategic issue is the problem of depot location, and a particular feature of the number of depots in a distribution network is their effect on stock holding. In cases where 80–90% of total costs are independent of total mileage, it is more profitable to examine means for reducing fixed costs (including the size of the delivery fleet) than to invest an inordinate amount of resources on vehicle scheduling. However, the level of sophistication of distribution studies is likely to grow, as management becomes increasingly concerned about rising costs.

> "Providence never intended to make the Management of publick Affairs a Mystery, to be comprehended by a few persons of sublime Genius, of which there are three born in an age".
>
> Swift (Gulliver's Travels)

INTRODUCTION

THE PHYSICAL distribution of goods from their point of manufacture to the ultimate users has always been at the heart of industrial development, and its central role in international trade is well recognised. Even though the distribution function is often thought to lack the glamour associated with marketing, and to a lesser extent perhaps even production, its possible effect on the performance and financial well-being of a manufacturing enterprise can be so far-reaching, that it is not surprising that distribution activities are increasingly subjected to a searching scrutiny. As for companies specifically engaged in the distribution business, keener competition and mounting costs have made management painfully aware of the need for greater sophistication and analysis.

The sums of money involved are certainly not insignificant; although estimates of the total cost of distribution naturally depend on various definitions and assumptions, one estimate put the cost in 1974 in the UK in the order of magnitude of £4000m, or about 6% of GNP, and against such figures even small improvements in percentage terms can involve large sums of money. It is not unreasonable therefore to expect a greater amount of managerial resources than hitherto to be devoted to the study of various aspects of distribution problems.

These problems can be broadly delineated at three levels, as suggested by the following examples:

Strategic level

for the long term—market evaluation, profitability analysis, pricing policy, diversification vs specialisation, working conditions, number and relative size of depots in the system, large capital expenditure, size of vehicles, inventory control, ordering and dispatch procedures, the use of computers.

Tactical level

for the short term—organisation and layout of depots, price quotation, vehicle replacement, maintenance procedures, handling equipment, manpower scheduling, fixed routes planning, debtors control.

Operational level

for day-to-day activities—route scheduling, handling breakdowns and priorities, queries and complaints, invoicing, general expediting.

Vol. 5
1977

There is nothing hard and fast about these categories, and clearly some problems that in one company may be classed as strategic in nature would be regarded as tactical in another. What is important to point out is that decisions are made at uneven frequencies, some (such as determining the number of depots in a distribution network) may be appropriate to review every few years, whereas others (such as vehicle routing) need to be undertaken on a daily basis, yet these decisions interact in the way they affect the total cost of the system. In any analysis of performance it is therefore necessary first to define the boundaries of the system in question and secondly to ensure that sub-optimisation of one of its components is not achieved to the detriment of the system as a whole.

Figure 1, for example, describes a logistics system which starts with a manufacturing plant and involves transportation to depots (called 'trunking'), storage at the depots, and delivery to customers [4]. The total cost = the cost of the depots + the trunking costs from the plant to the depots + the delivery costs from the depots to the customers. The definition of the system in Fig. 1 may not always be appropriate (it ignores, for example, the supply of raw materials to the plant, but even within this system it is possible to minimise the cost of trunking by having one depot located next to the plant at the expense

of a dramatic increase in delivery costs to customers; conversely, delivery costs can be reduced by having a very large number of depots, since delivery costs depend on the distance from depot to customer, but the trunking costs and the costs of the depots themselves would inevitably rise. The problem is therefore to find the number and location of depots that would minimise the total cost function. This approach (which is described in some detail elsewhere [4]) presupposes that there are two fundamental problems that need to be examined, the first is the strategic problem of depot location (and this term has come to include by implication the number of depots) and the second is the tactical problem of vehicle routing, so that optimisation is performed in tandem: for a given set of depots and a given configuration of customers the best routes are computed to provide an expected delivery cost, and the process is repeated in a systematic fashion for various possible solutions to the depot location problem, until no further improvement in the total cost function is obtained.

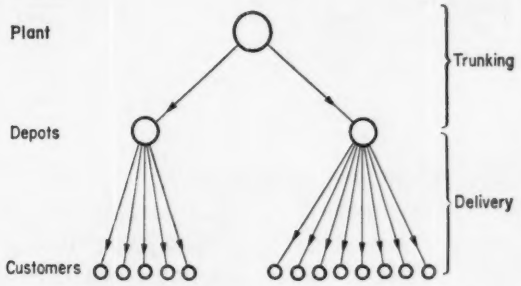


FIG. 1. *A distribution system* [4].

There is a third fundamental problem that impinges on the cost function and on the pattern of delivery, namely customer service. Not only does it affect inventory levels, and hence warehousing costs, but also the utilisation of vehicles and delivery routes. A further reference to the impact of customer service is given later, but the point needs to be made at the outset that the investigator of any business enterprise—and this observation equally applies to distribution systems—must be constantly aware of the interdependence of many of the variables that impinge on his study, and he must therefore exercise due care in the way that he isolates some of these variables and in the assumptions that are inherent in his methodology.

We first turn to the means by which the overall performance of an enterprise can be examined. It is not suggested, of course, that the performance criteria discussed here are the only ones that management is concerned with, but apart from figuring prominently in most managerial discussions, they serve to illustrate the method by which these criteria and others can be analysed.

THE RATE OF RETURN

A useful way of analysing the performance of an industrial enterprise is to start with a key measure, such as the rate of return on investment, and attempt to determine the possible effects that changes in various parameters are likely to have on this key measure. In this way not only is it possible to identify the major variables that the rate of return is sensitive to, but management may actively seek to change these variables in a particular way in order to achieve desirable improvement in performance.

The profitability model that I refer to, and the examination of its components, were originally proposed for a manufacturing enterprise [2], but there is no reason why precisely the same analysis could not apply to a distribution company, or to the distribution function within a company. The rate of return model is given by

$$\text{Rate of return} = \frac{\text{profit}}{\text{total investment}} = \frac{\text{unit profit} \times \text{volume}}{\text{total investment}}$$

or

$$r = \frac{(p - c)V}{I} \tag{1}$$

where

r = rate of return = ratio of profit (before tax) to total investment;
p = unit price = average rate charged per unit delivered;
c = unit cost = average cost of processing a unit through the system;
V = volume of sales, or number of units delivered to customers;
I = total investment = fixed assets + working capital.

Now, let us suppose that the major components that make up the rate of return r are subject to change, so that

p changes to $p(1+p^*)$;
c changes to $c(1+c^*)$;
V changes to $V(1+V^*)$

and these changes will cause r to change to $r(1+r^*)$. If it is assumed that the total investment I remains unchanged, it is then possible to determine the magnitude of the relative change r^*, which, ignoring second order effects, can be shown to be [2]

$$r^* = \frac{p^*}{a} - \frac{1-a}{a}c^* + V^* \tag{2}$$

where a is the profit margin $(p-c)/p$. This expression allows us to assess the effect on r^* of given changes in the system for a given value of a. For example, if the unit price increases by 5% and the volume decreases by 8%, and if the unit cost remains unchanged, then r^* (in %) becomes $(5/a) - 8$. The role of the

Vol. 5
1977

profit margin α here is quite crucial, the smaller it is the greater is the impact of change on the rate of return; for example if $\alpha = 10\%$ then the rate of return will improve dramatically by about 42%, whereas if $\alpha = 20\%$ the improvement will be only about 17%.

The above expression for r^* is an approximation, based on two assumptions, first that the total investment is unchanged and secondly that second order effects (for example terms which include p^*V^*) are ignored. It is, of course, a relatively simple exercise to develop a more elaborate expression without resort to such assumptions. Furthermore, it is possible to distinguish between fixed assets and working capital in our consideration of total investment and to incorporate into the analysis such factors as the price elasticity of demand, the effect of turnover or total costs on working capital requirements, as well as the impact of interest rates on borrowings, and details of such analyses are given elsewhere [3].

A UNIT COST MODEL

Similarly, and in conjunction with the model for the rate of return, it is possible to construct a model for the unit cost and determine the effect of changes in its components on cost performance. The simplest unit cost model is based on a distinction between fixed costs and variable costs, so that the cost per unit of volume throughput is [2]

$$c = s + \frac{F}{V} \tag{3}$$

where

s = variable cost per unit;
F = fixed costs;
V = volume.

Simple calculations soon reveal that the proportional costs have a profound effect on the way that a change in any of the cost components can affect the total cost. The incremental analysis yields the following result

$$c^* = f_1 s^* + \frac{f_2}{1 + V^*}(F^* - V^*) \tag{4}$$

where

$f_1 = s/c$ = proportion of the variable cost in the total unit cost
$f_2 = (F/V)/c$ = proportion of the fixed cost element in the total unit cost

and

$$f_1 + f_2 = 1.$$

The equation for c^* is fundamental to understanding the reasons for the relative effects of various components in the cost model. In the case of distribution the fixed costs account for a predominant proportion of the total, so that if for example $f_1=0{\cdot}2$ and $f_2=0{\cdot}8$ then the effects on the unit cost would be as shown in Table 1.

TABLE 1. EFFECTS ON UNIT COST

Change (in %)			Effect (in %)
s^*	F^*	V^*	c^*
−10	0	0	−2
0	−10	0	−8
0	0	+10	−7·3
−10	−10	+10	−16·5
0	0	−10	+ 8·9

We can see that the effect of reducing fixed costs greatly outweighs the effect of reducing variable cost, that the effect of increasing the volume is also quite substantial, and that the combined effect of all three changes is not simply the sum of the individual effects.

Three important observations need to be made about these models. The first is that the effect of changes in parameters need not be symmetrical; for example, in Table 1 an increase in V by 10% results in a decline in unit cost of 7·3%, whereas a reduction V by 10% results in the unit cost rising by 8·9%.

Secondly, the unit cost model is constructed as a simple sum of unit fixed costs and variable costs. It is, of course, possible to examine a much more elaborate model, for example by substituting

$F =$ F_1 (warehousing, fixed costs of the fleet, administration)

$+F_2$ (inventory holding costs, reflected by the level of customer service)

$+F_3$ (interest charges on borrowing, other than those accounted for in the inventory costs).

In this way, the effect of each of the elements of fixed costs can be explicitly ascertained by following the same methodology of incremental analysis, except that F^* in (4) needs to be expressed in terms of its individual components.

Thirdly, the examples in Table 1 suggest that changes in each of the components can occur independently of each other, whereas in reality any given change may trigger off a sequence of changes in the other components and each change will have its own effect on the unit cost and on the rate of return. Take, for example, the problem of customer service: a decline in the number of days which elapse in meeting customer orders may well affect the total volume, and a

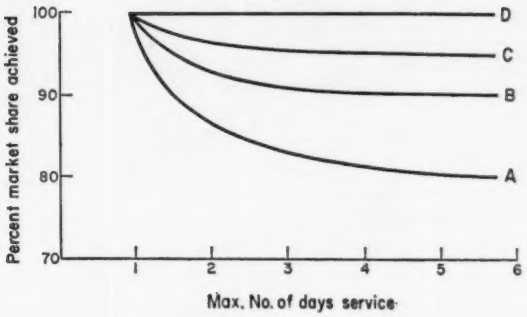


FIG. 2. *Market share affected by service level (Shycon* [10]).

typical set of alternative assumptions proposed by Shycon [10] is shown in Fig. 2 ranging from alternative A which depicts a loss in volume of up to 80% when the delivery lead time increases (from 1 day to 6) to alternative D which assumes no loss at all. Each of these alternatives can then be considered in turn to determine its consequences on the number of vehicles in the fleet and on the other warehousing fixed costs, as well as on the variable costs per unit, and all these changes can then be taken into account in estimating the total distribution costs and net profit as shown in the example in Table 2, quoted from a case study by Shycon [10] and based on curve A in Fig. 2. The example, alas, is not explicit in stating how the various cost figures have been arrived at (e.g. it is not clear why a decline of 20% in sales volume should require a reduction in

TABLE 2. AN EXAMPLE OF COSTS AND PROFIT
(Shycon [10])

Max. del. time/days	1	2	3	4	5	6
% market potential (curve A, Fig. 2)	100	87	83	81	80·5	80
Dollar sales volume ($m)	160	139·2	132·8	129·6	128·8	128
Dollar profit contribution at 30% ($m)	48	41·76	39·84	38·88	38·64	38·4
Number warehouses	50	20	14	10	8	7
Distribution cost ($'000)						
Transportation	9600	7516	7012	6739	6646	6579
Delivery	2909	5568	6989	8640	9908	10,667
Warehousing	3762	2187	1820	1560	1422	1349
Inventory at 20%	10,050	6749	5881	5240	4889	4694
Total distribution cost	26,321	22,020	21,702	22,179	22,865	23,289
Net contribution ($'000)	21,679	19,740	18,138	16,701	15,775	15,111

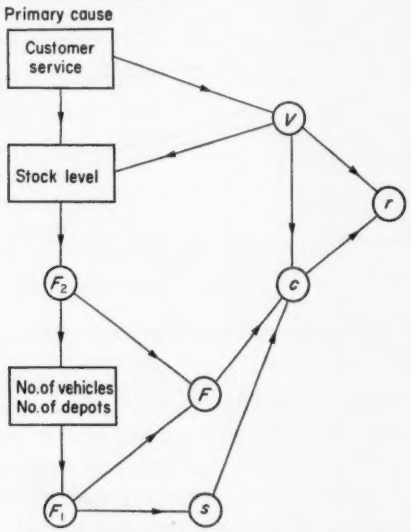


FIG. 3. *An example of a chain reaction to a change in customer service.*

Vol. 5
1977

the number of depots from 50 to 7, nor why elimination of 86% of the depots should correspond to a 64% reduction in warehousing costs), but it serves to illustrate a simple methodology for computing certain key cost components and the resultant profit. The incremental analysis that I am advocating is more detailed in concept, in that it seeks to determine the sequence of effects caused by a primary change, as graphically shown in the example in Fig. 3. Such analyses should become an integral part of the planning function and should be systematically carried out.

ON CUSTOMER SERVICE

It would be wrong, of course, to suggest that profit, cost, or the rate of return are the only measures of performance that need to be monitored in evaluating a distribution system; they are the most commonly accepted criteria for overall financial performance, but other measures need to be devised for particular aspects of the operation. Customer service, for example, is not confined to the time lag in delivery (Fig. 2) but encompasses many other issues, such as:

1. information (on stock availability, delivery dates, invoices);
2. dispatch (time lag for normal orders, emergency supply);
3. physical delivery (packaging, handling facilities for unloading, breakage, paper work);

4. stock control (cost of processing orders, incidence of stock-outs);
5. payment (invoicing, credit facilities);
6. complaints.

Some of these issues readily lend themselves to quantitative measures, as for example the distribution of time lags for delivery, the percentage of deliveries completed on time, the frequency of stockouts, the number of complaints; others have to make do with more descriptive and subjective assessments. Thus, in attempting to determine the possible effects of customer service in a scheme such as in Fig. 3, a change in any of these factors may account for a series of financial consequences.

Even simple clerical procedures may affect customer service, as illustrated by the following example: when orders for multi-items arrive at a store, they can either be processed through stock control first (to debit the stock cards by the amounts to be issued) and then through the store to pick the items for dispatch, or they can be sent direct to the storeman for picking and then to stock control for updating the records. The former procedure ensures that the stock cards take account of the orders in progress, so that inquiries about stock availability are answered with a high degree of accuracy, but actual dispatch to customers is thereby delayed (in one case that came to my notice this delay accounted for at least 2 days); the latter method puts priority on dispatch, which is good for delivery service, but results in the stock records being permanently out of date.

DEPOTS IN A DISTRIBUTION SYSTEM

I would like to turn now to the two problems that dominate the literature on distribution, namely depot location and vehicle scheduling. My object here is not to review the vast technical literature on these two topics, but to make some comments on applications.

The methodology for solving the depot location problem is well documented and is capable of handling a wide variety of complications. For example, one of the major complications often encountered is the discontinuities in haulage rates for trunking from the factory to depots, when rates are more dependent on destination than on distance, but for the purpose of determining a practical solution and conducting sensitivity analyses such complications do not present serious problems. As an example I refer to the case study described in [4] of an engineering company seeking to determine the optimal number of depots for distribution to 37 customer areas in Wales (Fig. 4), where the solution to have one depot or none (in addition to delivery from the factory direct) depended on the local delivery cost (Fig. 5). The effect of this cost led to a sensitivity analysis, given in Fig. 6, which clearly showed the range of local delivery costs for which any given solution remained dominant (in the example shown one of the four

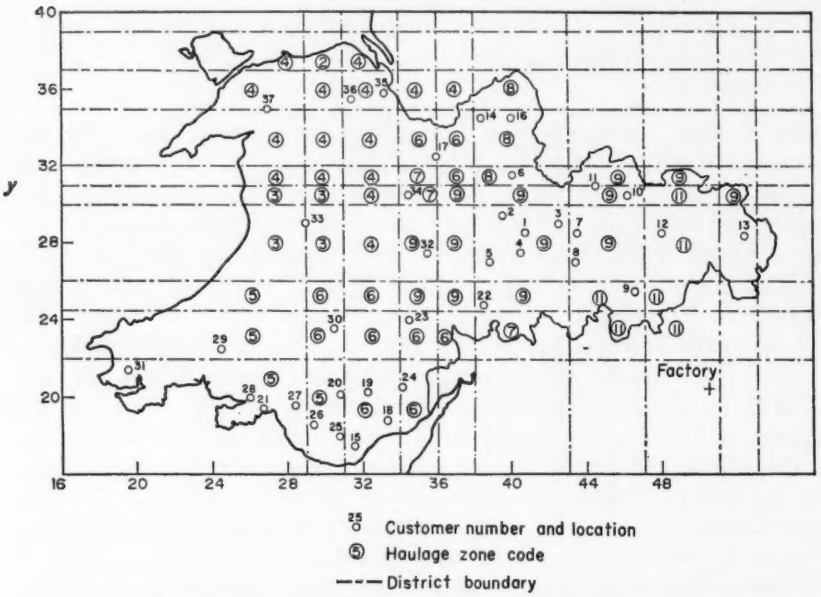


FIG. 4. *Map of area showing customer locations and haulage zones* [4].

solutions was dominated throughout). Another type of sensitivity analysis is provided by an iso-cost map, as shown in another example in Fig. 7, and by the cross-sections through the cost domain in Fig. 8.

What is not sufficiently realised is the effect that the number of depots in the system has on stock holding. Take the case of two depots, which have standard deviations of demand of σ_1 and σ_2 respectively. If the safety stock policy is the same in both depots, the safety stock for the two depots would be $k\sigma_1$ and $k\sigma_2$ respectively, where k is a constant that determines the percentage of customer demand that can be met without stockouts. Thus, the total safety stock in the two depot system is

$$S(2) = k\ (\sigma_1 + \sigma_2).$$

Now, if the two depots are combined into one to serve the same demand pattern, the standard deviation of demand for the combined depot is $\sigma = (\sigma_1^2 + \sigma_2^2)^{\frac{1}{2}}$ so that the safety stock (for the same stockout policy) is

$$S(1) = k\sigma = k(\sigma_1^2 + \sigma_2^2)^{\frac{1}{2}}$$

and the ratio between the two cases

$$\frac{S(1)}{S(2)} = \left[1 - \frac{2\frac{\sigma_2}{\sigma_1}}{\left(1 + \frac{\sigma_2}{\sigma_1}\right)^2}\right]^{\frac{1}{2}}$$

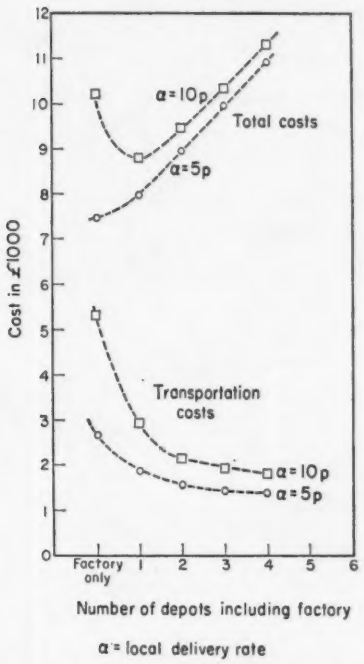


FIG. 5. *Total distribution costs and transportation costs* [4].

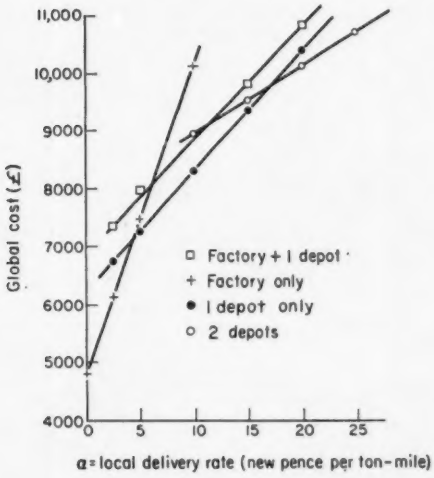


FIG. 6. *Sensitivity of the optimum solution to changes in local delivery rate* [4].

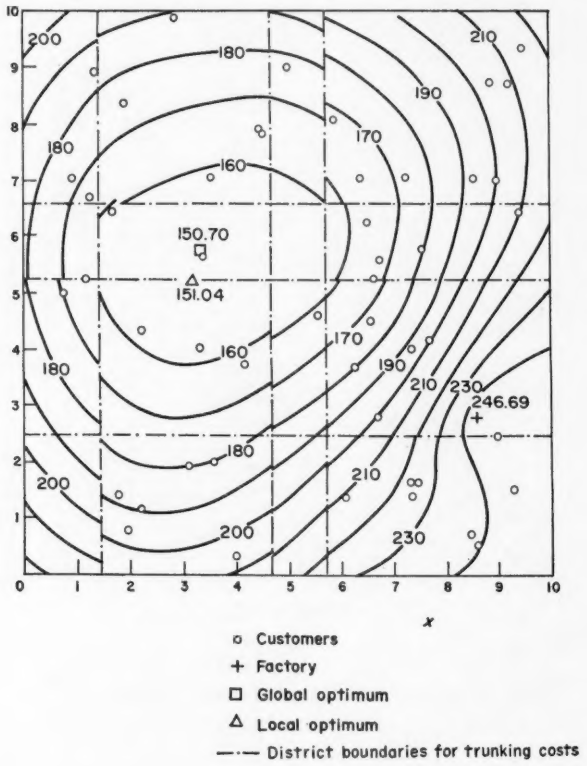


FIG. 7. *Iso-cost curves for one depot and one factory when the trunking cost is a function of distance and district* [4].

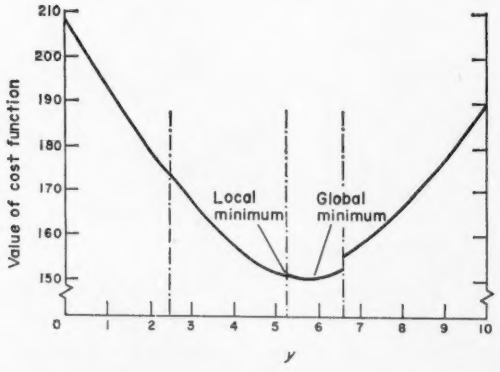


FIG. 8. *Total cost functioning* [4] (*section through* x = 3.40, *see Fig.* 7).

Vol. 5
1977

is shown in Fig. 9. For the special case where $\sigma_1 = \sigma_2$ we see that the safety stock in one depot is almost 30% below that required for two depots. Similarly, if m depots with a total safety stock $S(m)$ are combined into one, the ratio of the safety stocks is

$$\frac{S(1)}{S(m)} = \frac{(\sigma_1^2 + \sigma_2^2 + \ldots)^{\frac{1}{2}}}{\sigma_1 + \sigma_2 + \ldots}$$

and for the special case $\sigma_1 = \sigma_2 = \ldots$ the ratio is

$$\frac{S(1)}{S(m)} = \frac{1}{\sqrt{m}}$$

as shown in Fig. 10. Thus, a reduction in the number of depots from $m+1$ to m has the effect that

$$\frac{S(m)}{S(m+1)} = \sqrt{\frac{m}{m+1}}$$

and the consequent reduction in the safety stock is particularly noticeable when m is small (which then corresponds to a relatively large reduction in the number of depots in the system), as shown in Table 3.

TABLE 3. REDUCTION IN SAFETY STOCK FOR m DEPOTS COMPARED WITH $m+1$

m	1	2	3	4	5	8	10	20
Reduction in safety stock (%)	29·3	18·4	13·4	10·6	8·7	5·7	4·7	2·4

ON VEHICLE SCHEDULING

Of the two major problems that dominate the distribution literature, depot location and vehicle scheduling, it is the latter that has attracted the attention of numerous research workers, particularly over the last ten years. The fascination of the problem lies in the fact that it is embedded in some classical OR (operational research) problems: TSP (the travelling salesman problem to minimise the tour through a given number of customers), MTSP (the multi-TSP, where several salesmen start and finish at a given point, the depot), CTSP (the chromatic TSP, to minimise the number of salesmen when each is subject to a maximum distance constraint), and VSP (the vehicle scheduling problem, where capacity constraints need to be observed).

Since there is no efficient algorithm for finding the optimal solution to TSP, much of the effort of investigators in the field has been directed to developing heuristic solution methods with a view to improving the quality of the solutions

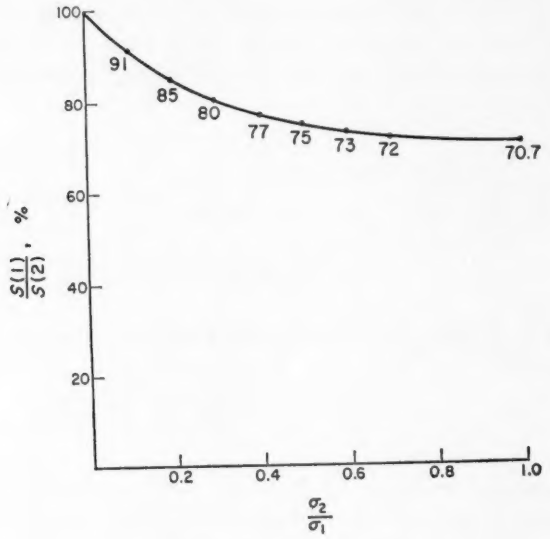


FIG. 9. *Safety stock: One depot vs two.*

Vol. 5
1977

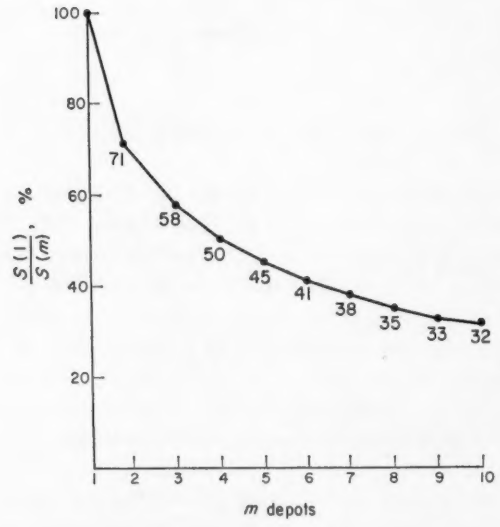


FIG. 10. *Safety stock: One depot replacing* m *depots.*

obtained and reducing the amount of computing time needed for the purpose. One method, for example, is based on the *r*-optimal tour [4], which starts with an arbitrary tour, then *r* links are removed and the disconnected chains are then reconnected to produce a shorter tour; the tour which cannot be further improved upon through this process is said to be *r*-optimal. The larger the value of *r* the better the quality of the solution, but computing time also rises very sharply. Experimental work has suggested that the 2-optimal and the 3-

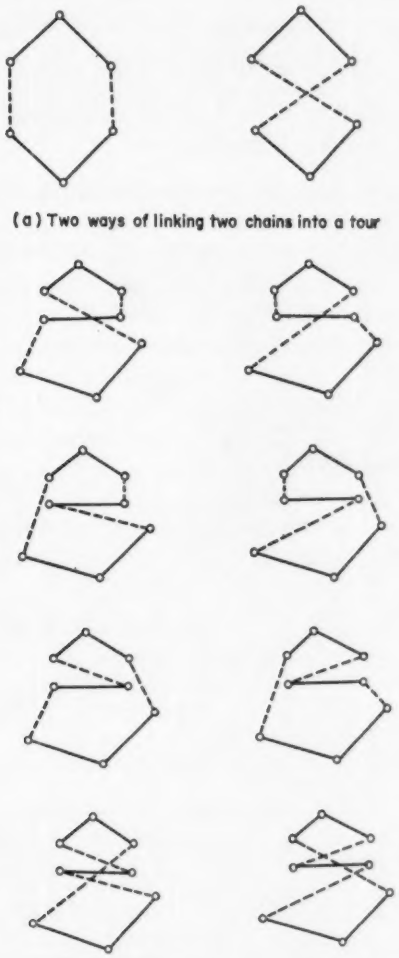


FIG. 11. *The 2-optimal and 3-optimal method* [4]. (a) *Two ways of linking two chains into a tour.* (b) *Eight ways of linking three chains into a tour.*

optimal methods (see Fig. 11) generally produce very good results, the former being very fast computationally.

There is no doubt that these methods can be improved upon, but I believe that much further work aimed at reducing tour lengths is likely to be of a marginal practical benefit. The reason for that lies in the simple fact that a reduction in the total distance travelled does not in general affect the fixed costs of a fleet, and these costs are predominant. In one documented study [5] an attempt was made to construct a cost function per vehicle in a fleet of 234 vehicles with the various cost items grouped as follows:

Fixed costs = taxes + insurance + wages and expenses + depreciation.

Variable costs = fuel + oil + tyres + maintenance (= routine + non-routine) + genuine overtime.

A statistical analysis to relate the variable costs to the mileage travelled, to the number of 'drops' (i.e. customers visited) and to the tons carried, revealed that little was gained by the latter two variables (drops and tons carried), that the costs of oil and fuel were highly correlated with mileage and that the costs of tyres and maintenance (a breakdown of which is shown in Fig. 12) had a low correlation with mileage. A reasonable model for the total cost per vehicle was found of the form $a + b \times$ (mileage), where the fixed costs a constituted 83% of the total. In two other cases for 12 and 20 vehicles respectively [5] the fixed costs amounted to 83–87% of the total before the fuel crisis, in a third case in 1974 a fleet of 45 vehicles covering over 20,000 miles a week incurred under 15% of the costs in fuel and oil.

Quite clearly, then, any effort invested in cost control should concentrate more on a reduction in the size of the fleet than on a reduction in mileage. For

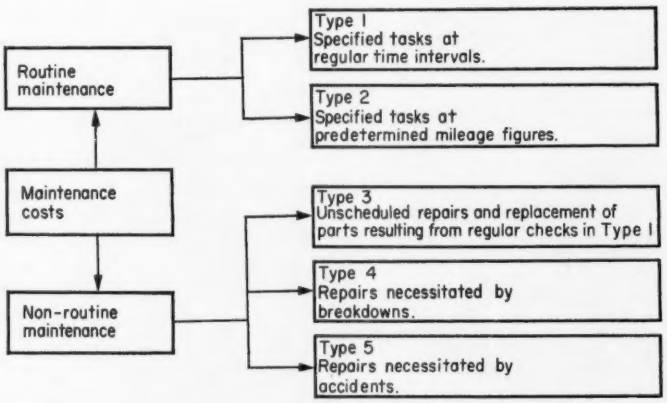


FIG. 12. *A breakdown of maintenance costs* [5].

Vol. 5
1977

a fleet with 80% fixed costs a saving of 5% in mileage as a result of sophisticated route planning (and this is a high figure, if it purports to achieve economies over and above the performance of relatively simple methods, based for example on 2-optimal tours) then the resultant effect on the total cost would be only 1%, whereas a similar reduction in fleet size would be far more significant, even if the total mileage is thereby somewhat increased.

I, therefore, favour the idea of simple heuristics that aim at achieving high levels of vehicle utilisation. From analytical and computational work [4, chapter 8] it is possible to estimate the tour length to visit n customers; for example, if n are points randomly distributed in an area A, the expected tour length is

$$L(n) = k\sqrt{nA} \tag{5}$$

(where $k=0{\cdot}75$ is regarded as a reasonable estimate), and when the depot is situated at some distance from the centre of a 10×10 square in which n customers are to be included in a vehicle tour, the expected tour length can be obtained from computational results such as in Fig. 13 [4]. Thus, the number of customers that can be allocated to a vehicle within a given distance constraint can be estimated, and then a procedure like the *sweep algorithm* can be used. This method relies on a radial scan (see Fig. 14), like a scan on a radar screen, to identify the customers that need to be supplied, and when the requisite

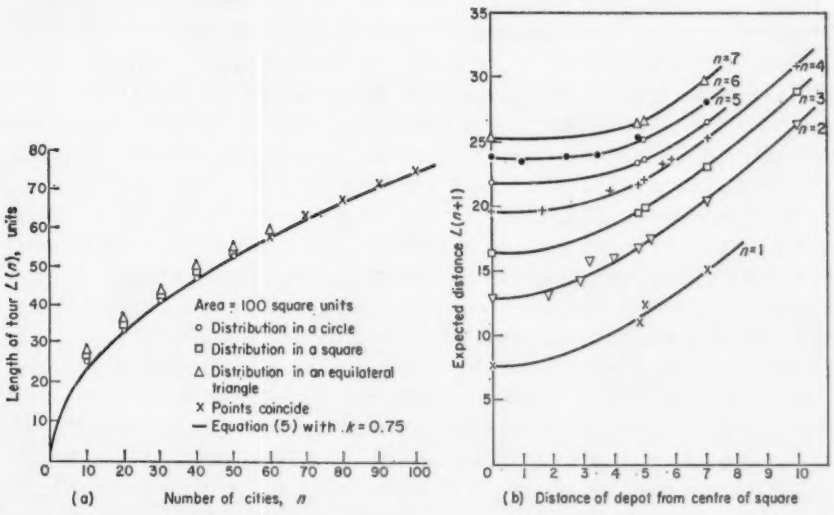


FIG. 13. *Expected tour lengths* [4]. (a) *Expected travelling salesman tour lengths.* (b) *Expected length* (L(n+1)) *of a travelling salesman tour through* n *customers and one depot.*

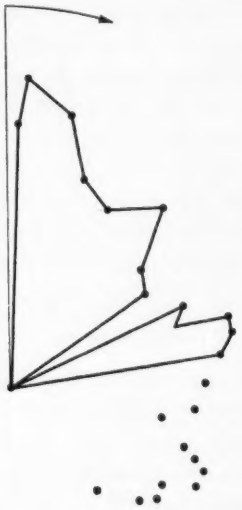


FIG. 14. *The sweep algorithm.*

Vol. 5
1977

number that complies with the maximum total distance constraint is reached a tour can be drawn in the sequence in which the points emerge from the radial sweep. The resultant tour is not necessarily optimal, as shown in the example in Fig. 15, where a distant point from the depot is followed by a near point, which is then followed by a distant point. If the tour is then subjected to a 2-optimal procedure (which is extremely fast for the relatively small number of points involved), better solutions can quickly be found.

In cases where the distance constraint is the main cause for low vehicle utilisation, the sweep algorithm can be modified by the provision of an inner zone, which is excluded from the outer zone scanning process, i.e. the main radial sweep identifies the customers in the outer zone that should be allocated to a given vehicle route, and then customers can be added in the inner zone to fill up the vehicle capacity. Thus, the outer zone caters for the distance constraint and the inner zone deals with the capacity constraint. The 'topping up' by inner-zone customers has only a marginal effect on route lengths and may often allow a more frequent service to be offered within that zone. Such a 'dual sweep' method, with separate scanning in each zone (Fig. 16), may prove to be both effective and simple for vehicle scheduling in practice, and it is methods like this that need to be further investigated (note that the routes in Fig. 16 can be further improved if the 2-optimal routine is applied to the whole of each route and not just to its individual portions in the two zones).

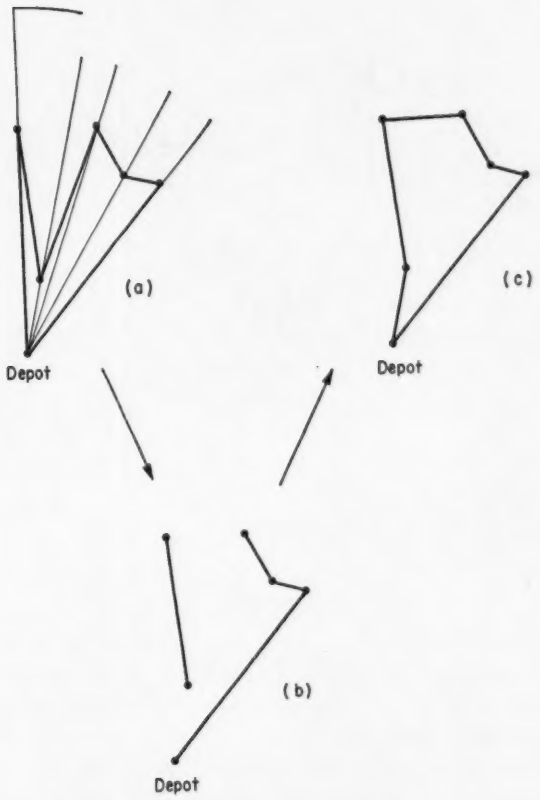


FIG. 15. *Improvement on the sweep route;*
(a) *The sweep.* (b) *2-Optimal routine.* (c) *Improved route.*

THE COMPUTER IN DISTRIBUTION

Finally, a few observations may be apposite on the role and possible impact of the computer on distribution management. It was suggested earlier that management tasks can conveniently be categorised as strategic, tactical and operational, and in each case the computer can evidently be used to store and retrieve information, as well as perform searching analyses for planning and control purposes. A flow-chart by Sutton [11] that attempts to describe the framework of an overall distribution study is shown in Fig. 17, starting with the information that needs to be assembled, and proceeding with sophisticated tools, such as linear and mixed integer programming, cost–benefit analysis and a PERT network for implementation. The flow-chart is evidently oriented

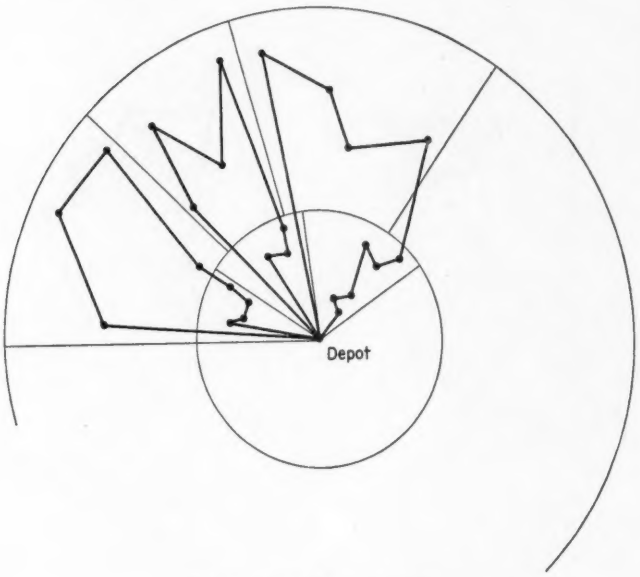


FIG. 16. *The dual sweep method.*

towards an *ad hoc* study (only parts of which need to be carried out with the assistance of a computer) of a given or a new distribution system and is not intended for operational purposes.

Where the computer can be of immense value is first in the storage and retrieval of information, for which an adequate data base is needed in order that management can readily get reliable answers to such questions as [1]:

What is the stock level of a given commodity in a given depot?

When is the next replenishment of stock going to be made?

Can a given order be met in a specified time?

What has happened to a particular order for which there is a query (about delivery, price or quantity)?

What is the credit rating of customer *X* and should he be given further credit?

Can an urgent call be added to already planned vehicles routes?

How often has vehicle *Y* broken down in the past six months?

With the advent of remote terminals and mini-computers there is no reason why the computer should not increasingly be used as an aid to the decision making process, even at the operational level (where until now the use of computers has been relatively limited, as for example in routing and dispatch).

An example of the level of cost detail which a computer can provide is cited

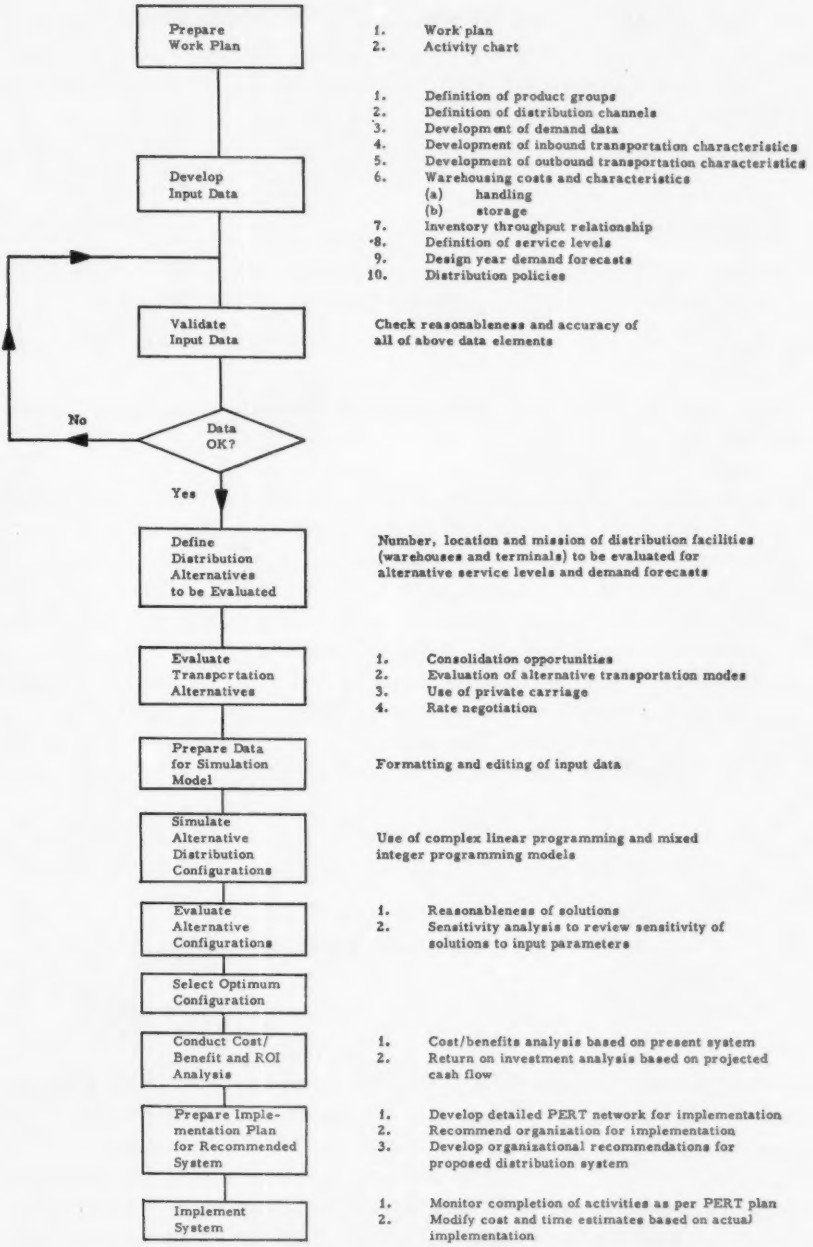

Prepare Work Plan
1. Work plan
2. Activity chart
Develop Input Data
1. Definition of product groups
2. Definition of distribution channels
3. Development of demand data
4. Development of inbound transportation characteristics
5. Development of outbound transportation characteristics
6. Warehousing costs and characteristics
(a) handling
(b) storage
7. Inventory throughput relationship
8. Definition of service levels
9. Design year demand forecasts
10. Distribution policies
Validate Input Data
Check reasonableness and accuracy of all of above data elements
No
Data OK?
Yes
Define Distribution Alternatives to be Evaluated
Number, location and mission of distribution facilities (warehouses and terminals) to be evaluated for alternative service levels and demand forecasts
Evaluate Transportation Alternatives
1. Consolidation opportunities
2. Evaluation of alternative transportation modes
3. Use of private carriage
4. Rate negotiation
Prepare Data for Simulation Model
Formatting and editing of input data
Simulate Alternative Distribution Configurations
Use of complex linear programming and mixed integer programming models
Evaluate Alternative Configurations
1. Reasonableness of solutions
2. Sensitivity analysis to review sensitivity of solutions to input parameters
Select Optimum Configuration
Conduct Cost/ Benefit and ROI Analysis
1. Cost/benefits analysis based on present system
2. Return on investment analysis based on projected cash flow
Prepare Implementation Plan for Recommended System
1. Develop detailed PERT network for implementation
2. Recommend organization for implementation
3. Develop organizational recommendations for proposed distribution system
Implement System
1. Monitor completion of activities as per PERT plan
2. Modify cost and time estimates based on actual implementation

FIG. 17. *Study methodology for a distribution system (after RM Sutton* [11]).

by Helms [6]. He describes the operations of the Martin-Brower Co., which supplies 8000 restaurants with 2000 items, using a fleet of 200 tractors and 300 refrigerated and dry trailers. The depot costs are divided into the following categories:

inbound handling;

order picking;

outbound handling;

storage (building, inventory insurance, rent and/or depreciation, rates, repairs and maintenance, interest, protection services and miscellaneous);

administrative (data and information processing, including orders and invoicing);

miscellaneous warehousing support (stock rotation, housekeeping, in-plant maintenance).

Whether such a detailed breakdown is desirable, and if so whether it is reliable, is open to question. The more cost categories there are, the more difficult it is to ascertain with any given degree of confidence the values to be attached to the various activities. In the example given by Helms the items were divided into three main types (two for dry foods and the third for frozen foods); the inbound handling costs per case was the same for all items in each type, but varied from type to type, whereas the outbound handling costs were the same throughout, and storage costs varied from item to item (Table 4). Clearly the assignment of any cost values depend here on accounting conventions and also on the availability of sufficient information to allow a differentiation between various items to be made.

Vol. 5
1977

Perhaps a smaller number of categories than that cited by Helms would be more appropriate, and it is interesting that his figures yield the following cost breakdown:

handling (inbound, outbound + warehousing support)	21·4%
picking	20·3
administration	29·3
storage (including interest)	28·9

However, it is not so much the level of detail which I wish to highlight here, as the fact that information of this kind, if properly coded, can provide a cost breakdown by items (or groups of items), by activities, by accounts, or by geographical demand regions. Not only can it become a basis for calculating manpower requirements, but it can serve as a focus for investigations into performance and the likely effects of proposed changes in the system. These problems are particularly important in the light of rising labour costs, an example for which is given by the trend in the US, cited by Lukenda [9], who

T.S 77

TABLE 4. A PRINTOUT FOR ITEM COST PER CASE ($)
(Given by Helms [6])

Product category	Item code	Outbound handling	Inbound handling	Order picking	Storage	Administrative	Misc.	Total
Dry food	7526	0·03596	0·02584	0·14495	0·03531	0·13123	0·03051	0·40380
	7527	0·03596	0·02079	0·14495	0·20064	0·13123	0·03051	0·56407
	7529	0·03596	0·02079	0·14495	0·21666	0·13123	0·03051	0·58009
	7530	0·03596	0·02079	0·14495	0·18917	0·13123	0·03051	0·55261
	7531	0·03596	0·02079	0·14495	0·53685	0·13123	0·03051	0·90028
	7557	0·03596	0·02079	0·14495	0·02031	0·13123	0·03051	0·38374
	7558	0·03596	0·02079	0·14495	0·14879	0·13123	0·03051	0·51222
	7564	0·03596	0·02584	0·14495	0·16149	0·13123	0·03051	0·52998
	7565	0·03596	0·02079	0·14495	0·09070	0·13123	0·03051	0·45413
	7566	0·03596	0·02079	0·14495	0·69696	0·13123	0·03051	1·06039
	7567	0·03596	0·02584	0·14495	0·02996	0·13123	0·03051	0·39845
	7568	0·03596	0·02079	0·14495	0·06428	0·13123	0·03051	0·42772
	7569	0·03596	0·02584	0·14495	2·03727	0·13123	0·03051	2·40576
	7570	0·03596	0·02079	0·14495	0·07083	0·13123	0·03051	0·43426
	7571	0·03596	0·02079	0·14495	0·40128	0·13123	0·03051	0·76471
	7572	0·03596	0·02079	0·14495	0·18651	0·13123	0·03051	0·54994
	7573	0·03596	0·02584	0·14495	0·11616	0·13123	0·03051	0·48465
Frozen food	7517	0·03596	0·02584	0·05390	0·11187	0·13123	0·03051	0·38931
	7519	0·03596	0·02584	0·05390	0·27387	0·13123	0·03051	0·55131
	7520	0·03596	0·02584	0·05390	0·19353	0·13123	0·03051	0·47097
	7521	0·03596	0·02584	0·05390	0·09809	0·13123	0·03051	0·37553
	7522	0·03596	0·02584	0·05390	0·15887	0·13123	0·03051	0·43631
	7523	0·03596	0·02584	0·05390	0·10231	0·13123	0·03051	0·37975
	7524	0·03596	0·02584	0·05390	0·09882	0·13123	0·03051	0·37627
	7563	0·03596	0·02584	0·05390	0·62534	0·13123	0·03051	0·90278
French fries	7481	0·03596	0·04608	0·05390	0·07852	0·13123	0·03051	0·37619
Account total		0·03596	0·02951	0·09070	0·12942	0·13123	0·03051	0·44732

also pointed out that between 1972 and 1975 the labour cost increase outstripped the cost increase of fork lift trucks (see Tables 5 and 6), thereby suggesting that an increased level of mechanisation in depots is bound to remain a topic for consideration by management. There should in principle be no difficulty in extending the data-base to include cost items on transportation and various characteristics of operations of the delivery fleet.

TABLE 5. WAREHOUSE LABOUR COST (UNION SHOP)
(JR Lukenda [9])

Year	Hourly cost	Annual cost	Percentage increase from Previous yr.	Base year
1970	$5·39	$11,211	—	—
1971	6·12	12,729	13·5	13·5
1972	7·02	14,601	14·7	30·2
1973	7·83	16,286	11·5	45·3
1974	8·16	16,973	4·2	51·4
1975	9·08	18,886	11·3	68·5
1976	10·13	21,070	11·6	87·9
Est. 1980	16·08	33,446		198·3

TABLE 6. LABOUR VS EQUIPMENT COSTS 1972 VS 1975
(JR Lukenda [9])

Year	Labour (Cost/hour)	Equipment (Lift truck)
1972	$7·02	$12,500
1975	9·08	14,400
Increase	$2·06	$1900
	or 29·3%	or 15·2%

It is interesting in this context to quote the 1972 survey of LaLonde and Auker [8] on the applications of computers in distribution and an extract is given in Table 7. The survey covered 107 firms (mainly members of NCPDM[1]) and the results reflect the predominance of conventional data-processing to do with inventory control and processing of orders. But judging from work which was then regarded as being at the development and planning stage, a higher level of sophistication is on the way, though surprisingly enough not in costing and budgeting and not in the use of simulation.

[1] National Council of Physical Distribution Management (in the US).

Vol. 5
1977

TABLE 7. EXTRACT FROM A SURVEY ON THE USE OF COMPUTERS
(LaLonde and Auker [8])

	In use (%)	Under development/planning (%)
Inventory control	84·1	8·4
Order processing	79·4	13·0
Bill of lading preparation	49·5	21·5
Shipment forecasting	41·1	24·3
Vehicle routing	11·2	19·7
Distribution costing and budgeting	6·5	0·9
Simulation	4·6	1·9

TABLE 8. DIVISION OF TIME AMONG DISTRIBUTION RELATED ACTIVITIES—1975
(LaLonde [7])

Activity	Time 0%	1–10%	over 10%
Packaging	32	64	4
Warehousing	4	41	55
Traffic management	9	49	42
Order entry and processing	15	60	25
Warehouse location	18	70	12
Inventory management	14	50	36
International management	48	42	10
Product planning	55	38	7
Sales forecasting	49	48	3
General management	25	27	48
Other	73	13	14

That these and allied problems are uppermost in the minds of distribution managers is evident from another LaLonde survey in the US [7], based on 132 executives who replied to his questionnaire. Asked to apportion their time among various activities, the answers summarised in Table 8 indicate warehousing, traffic management, order processing, inventory control and general management as their main preoccupations.

CONCLUSION

The world of the distribution executive is becoming more complicated, with rising costs and customer service demanding an increasing level of sophistication in the analysis of distribution activities. It is not surprising to see, therefore, a substantial number of graduates, many even with postgraduate qualifications, in the distribution function in the US (see Table 9, from LaLonde [7]), and this

trend is bound to have some effect in the UK as well, where the allocation of resources to the study of distribution systems in some detail is well overdue.

TABLE 9. EDUCATIONAL ATTAINMENT FOR DISTRIBUTION EXECUTIVES IN THE US—1975 (LaLonde [7])

	Manager	Director	Vice-President
High school graduate	2	0	5
College—no degree	10	10	0
College graduate	42	48	48
Some graduate school	10	11	14
Master's degree	36	31	33
Total (%)	100	100	100

REFERENCES

1. EILON S (1971) The computer in distribution. *Int. J. phys. Distrib.* **2**(1), 14–16.
2. EILON S and PHOCAS-COSMETATOS G (1976) Analysis of profitability components. EURO II Congress on Operations Research, Stockholm.
3. EILON S and PHOCAS-COSMETATOS G (1977) A profitability model for tactical planning. The International Symposium on Extremal Methods and Systems Analysis, Austin, Texas; also to be published in *Omega* **5** (6).
4. EILON S, WATSON-GANDY, CDT and CHRISTOFIDES N (1971) *Distribution Management; Mathematical Modelling and Practical Analysis.* Griffin, London.
5. EILON S, WATSON-GANDY CDT and HEILBRONN A (1971) A vehicle fleet cost model. *Int. J. Phys. Distrib.* **1**(3), 126–132.
6. HELMS TJ (1975) Can a computer model provide you with the cost information necessary to improve your distribution operation? *Proc. NCPDM Meet.* Chicago, pp. 529–558.
7. LA LONDE BJ (1975) The distribution manager today. *Proc. NCPDM Meet.* Chicago, pp. 1–22.
8. LA LONDE BJ and AUKER K (1972) A survey of computer applications and practices in transportation and distribution. *Proc. NCPDM Meet.*, Chicago, pp. 227–246.
9. LUKENDA JR (1975) Labor vs capital intense material handling systems. *Proc. NCPDM Meet.*, Chicago, pp. 337–365.
10. SHYCON HN (1975) Distribution can generate added profit (or loss)! *Proc. NCPDM Meet.*, Chicago, pp. 251–267.
11. SUTTON RM (1975) The systems approach to transportation improvement. *Proc. NCPDM Meet.*, Chicago, pp. 187–198.

ADDRESS FOR CORRESPONDENCE: *Professor Samuel Eilon, Department of Management Science, Imperial College, Exhibition Road, London SW7 2BX, UK.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

Forecasting: A Hybrid Approach

NOEL D URI*

Federal Energy Administration, Washington

(*Received July* 1976; *in revised form November* 1976)

It has recently been shown that the Box–Jenkins approach to forecasting time series is superior to an econometric approach over a relatively short horizon. The results here support this contention. A combination of the two approaches, however, proves to be clearly superior to either one separately. By taking account of changes in economic and weather related variables in a time series model, improved forecasts are obtained.

1. 5
77

CAPACITY expansion plans of electric utilities are based on forecasts of peak demands. The long lead times required for planning physical facilities, the high cost of capital, and the necessity of maintaining a high service reliability all contribute to the importance of accurate peak demand forecasts in the intermediate-term (3–10 years) [9].

Because the critical nature of accurate forecasts, forecasting methodology is continually being refined. For example, recently there has been a plethora of econometric modeling approaches that are somewhat complicated and require a large amount of exogenous information. That econometric models are not the final answer to good forecasts is now a well-established fact [7]. Poor predictive performance may result from improper specification and errors of estimation.

An alternate approach, Box–Jenkins time series analysis has been shown to provide accurate forecasts in the intermediate-term [11]. The approach has been extensively criticized, however, because if the resultant forecasts are poor, one is at a loss to provide an explanation why since they have no basis in economic theory [7]. Furthermore, if one desires to explain the behavior of an economic system and not merely to provide forecasts, then a Box–Jenkins method is unacceptable. Since it is void of economic theory, it cannot be used to test hypotheses for economic phenomena.

This paper is a first step in taking the basic Box–Jenkins structure and integrating it into an econometric system. The approach adopted is to look at a specific example.

* The author is an economist with the Federal Energy Administration. The views expressed are those of the author and do not necessarily represent the policies of the Federal Energy Administration or the views of other Federal Energy Administration staff members. He would like to thank CWJ Granger and an anonymous referee for helpful suggestions.

I. A HYBRID BOX–JENKINS ECONOMETRIC MODEL

Let Z_t denote the peak system load of a utility at time t (monthly), and assume, following Uri [11], that

$$(1 - B)(1 - B^{12}) \log Z_t = (1 - \theta B)(1 - \eta B^{12}) a_t \tag{1}$$

where B is the backward shift operator (i.e. $B^\rho Z_t = Z_{t-\rho}$) and the a_t are independent deviations.

This model specification was adopted because of the characteristics inherent in the data. Specifically, there is a distinct seasonal pattern. An annual peak is reached in August. The annual trough occurs in April. Secondly, there is an upward tendency in the level of the series. Finally, the size of the seasonal effect increases with an increase in the level of the series indicating that the seasonal effect is multiplicative.[1]

Both first differencing and seasonal adjustment as performed here are steps towards prewhitening the data; that is, producing a series having a generally flatter spectrum than the original.[2] Once this is accomplished, the next step in the process is to specify a model relying on one-way causality whereby deviations from the adjusted series, the a_t, are dependent on a set of economic and weather related variables. One has to be careful, however, when dealing with multiple series not to introduce a feedback loop such that the original series has an effect on the set of economic and weather related variables. If this arises, an alternative approach is necessary [3]. This is not a problem here. The particular market conditions of electrical energy are such that one-way causality results. The reason is that prices are largely set by public regulation and governed by policies as to what is in the public's interest rather than by cost or supply considerations alone. Consequently, price may be regarded as predetermined, at any rate in the short run. Additionally, weather variations are clearly unaffected by system loads as are changes in income in the short run.

Assume, now, that the parameters θ and η are estimated using a Box–Jenkins procedure from a historical data set consisting of $T = 108$ monthly peak system load observations (9 years of 12 monthly observations). Resulting from this procedure will be a sequence of estimates of a_t, t = 14, 15,..., 108 which are prewhitened [1, p. 380] that, in general, will vary through time. In particular one can realistically hypothesize that the a_t are functions of the price or the change in the price of electrical energy in constant dollars, income or the change in income in constant dollars, temperature related variables and so on and specify a regression model accordingly.

[1] Seasonality can change and it is not easy to deal with this in the Box–Jenkins approach. It could be argued that it is better to seasonally adjust the data before estimating the Box–Jenkins model.

[2] Prewhitening the data simply removes the autocorrelation among the error terms.

Vol. 5
1977

With this objective, let **a** represent a vector of observations on a_t, and let X denote a set of economic and weather related variables. Next, it is postulated that

$$\mathbf{a} = X\alpha + \mu \tag{2}$$

where α is the appropriate vector of constants and μ is the stochastic term.

It is not likely that μ will be heteroscedastic since the prewhitening of a_t necessarily leads to a homoscedastic **a** and hence μ [1, p. 380]. Consequently, efficiency will be obtained by ordinary least squares without resorting to generalized least squares.

Thus, in equation (2), α should be estimated by

$$\hat{\alpha} = (X'X)^{-1}X'\mathbf{a}. \tag{3}$$

This estimate of the vector α allows one to forecast the residuals which, upon combining their forecast with the Box–Jenkins forecast, results in an improved forecast of the original series.

Given this structure, the objective is to implement it and indicate the improvement in forecasting accuracy over, say, a purely Box–Jenkins model or a pure econometric model. Before this is accomplished, however, a discussion of the data is in order.

II. THE DATA

The model is implemented at the utility level. Data on the monthly peak system load covering the period January 1961 to December 1969 were used in the estimation procedure. The system data were obtained from a utility operating in a major metropolitan center in the western United States.[3] Currently, the company serves in excess of two million customers with the commercial sector consuming almost as much electrical energy as the residential and industrial sectors combined. The company operates a mixture of oil–gas fired fossil fuel plants and hydroelectric plants. Its total installed capacity is in excess of 10 GW.

Data on the set of economic and weather related variables used in the estimation consist of monthly observations on price, income, and weather and were obtained from several sources. Average prices, representing total revenue divided by total sales for the utility, were obtained from the specific utility. The income data used were the monthly average of average weekly earnings compiled by the Bureau of Labor Statistics [2]. The ideal measure, gross domestic product, is just not available by utility service area much less on a monthly basis. The alternative adopted was the best available. While it is not exact, for the purpose at hand it is sufficient since it reflects changing economic conditions,

[3] In return for the data, firm anonymity was agreed upon.

structural changes in the work force[4], and the life style of the majority of residential consumers. Both the price and income variables were measured in constant 1970 dollars.

Several weather variables including average heating degree days, average cooling degree days[5], monthly average temperature, and daily maximum and minimum temperatures were available. Preliminary analyses showed the maximum temperature on the day the peak monthly system load occurred to be the best explanatory weather variable. Using the monthly peak demand might produce some difficulty in estimating temperature response in other applications. Severe temperature on a given day preceded by mild temperature will not result in as great a peak demand as several days of severe temperature in succession.

III. EMPIRICAL IMPLEMENTATION

Three separate models were estimated: (a) Box–Jenkins Model, (b) Econometric Model, and (c) A Hybrid Model. The functional form of each model is discussed in order.

Vol. 5
1977

(*a*) *Box–Jenkins model*

Since the size of the seasonal effect increases with an increase in the level of the series, logarithms of the original monthly peak observations were taken in order to fit a model with an additive seasonal effect. Once this was completed, the size of the seasonal effect appeared to be roughly constant and remarkably regular.

Letting the observed peak load at time t be denoted by Z_t and the transformed value by v_t where $V_t = \log_e Z_t$. The first stage of the Box–Jenkins procedure is to difference the series $\{V_t\}$ until a stationary series is obtained. As the series has a trend and a seasonal pattern completing one cycle every 12 observations, the sample autocorrelation functions of the series $(1 - B)(1 - B^{12})V_t$ were examined. The standard error of each of the autocorrelation coefficients would be approx. 0·10 if the series was random. It was observed that the coefficients at lags 1 and 12 are statistically significant at the 0·95 level. The substance of this significance is that it is reasonable to consider a lag of greater than 12 to be zero.

Given these results, the model (which is the same as (1))

$$(1 - B)(1 - B^{12})V_t = (1 - \theta B)(1 - \eta B^{12})a_t \tag{4}$$

was tentatively entertained to represent the trend and seasonal dependence of the series.

[4] For example, long-term increases in the proportion of part-time workers in retail trades reduces the workweek and hence reduces average weekly earnings.

[5] A degree day is a unit measuring the extent to which the outdoor mean (average of maximum and minimum) daily dry-bulb temperature falls below (in the case of heating) or rises above (in the case of cooling) a base of 65°F.

The estimates of the parameters θ and η were obtained. The estimates are given in Table 1.

TABLE 1. BOX–JENKINS PARAMETER ESTIMATES

Parameter	Estimate	95% confidence interval Lower limit	Upper limit
θ	0·6483	0·4861	0·8105
η	0·4926	0·2902	0·6951

Residual sum of squares = 0·0899.
93 Degrees of freedom.

The autocorrelations of the residuals exhibit no systematic pattern and are quite small in magnitude. The average of the residuals is approx. 0·00097 and the estimated standard error is 0·03108, strongly suggesting that the residuals have zero mean. Consequently, (4) is taken to be the appropriate model for the series V_t.

Refining on this, suppose it is assumed that the residuals are independent and normally distributed with mean zero and variance σ_a^2. It can then be shown that for large samples, the autocorrelations will be independent and normally distributed with mean zero and variance $1/\sqrt{T}$ where T is the number of observations [8, p 124]. In the present case $T = 95$, so that the standard deviation of the autocorrelations is approx. 0·10. None of 48 computed autocorrelations were significantly different from zero.

(*b*) *Econometric model*

In the short-run the demand for electrical energy by an individual is determined primarily by his stock of energy-using capital goods. It would not be inconsistent with other work to assume that this is the only factor influencing demand. In the present analysis, however, this assumption is rejected as being inconsistent with the accepted theory of demand. Instead, a less restraining assumption underlying a flow-adjustment model of demand as described by Houthakker and Taylor [4] is used. In these models, the stock of electrical energy-using appliances, equipment, and machinery is assumed to be fixed in the short-run and the utilization of it is assumed to be entirely a function of normal economic and weather influences. In this situation a classical adjustment model of demand can be used. In particular assume that there is an equilibrium demand q_t^* for electrical energy by consumers. This demand is made a function of average price, p_t[6], and a vector of other relevant economic and temperature

[6] Average price is used in deference to marginal price or some combination of marginal and average prices. Recent research at the individual utility level has shown average price to be a superior price measure to either some type of weighted class of marginal rates or a combination of marginal and average price [10].

variables, Ω_t. To elaborate on this concept of equilibrium, suppose that expectations of price and the other variables are static; that is, consumers expect current price and so on to persist indefinitely. Under an expectation of this kind, all changes in these variables are expected to be permanent. Given sufficient time to adjust, for every set of price and other relevant economic and temperature variables, only one equilibrium quantity, q_t^*, will be demanded. In general functional form,

$$q_t^* = f(p_t, \Omega_t). \tag{5}$$

This level of demand is reached only in conditions of long-run equilibrium. A very simple adjustment process is assumed whereby

$$\frac{q_t}{q_{(t-1)}} = \left(\frac{q}{q_{(t-1)}}\right)^{\nu} \tag{6}$$

for $0 < \nu < 1$. Hence, actual demand by a consuming sector in the current period, q_t, is given by

$$q_t = (q_t^*)^{\nu} (q_{(t-1)})^{(1-\nu)}. \tag{7}$$

Vol. 5
1977

Implicit in this model is the assumption that the utilization rates for all electrical energy-using types of capital (appliances, machinery, etc.) are the same regardless of vintage and other characteristics. Such an assumption is unsatisfactory. It would be far superior to incorporate the utilization of capital stock directly in the demand equation. However, data are lacking and pending their procurement, the flow adjustment model is adopted as the most viable specification.

A further insight into the model is obtained by looking at ν. It is expected that the value of this parameter is between 0 and 1. The proportion of the demand adjustment completed in the first period is given by $1-\nu$. If ν is close to 0, demand adjusts quickly to changes in the causal factors while as ν close to 1 implies a slower adjustment of demand.

If it is assumed that q_t^* is log-linear of the form

$$q_t^* = B p_t^{\beta_0} y_t^{\beta_1} w_t^{\beta_2} e^{\epsilon_t} \tag{8}$$

where q^* is the desired peak demand, p is average price, y is average weekly earnings, w is maximum daily temperature, ϵ is an error term, and β, β_0, β_1, β_2 are parameters; the estimating equation becomes

$$\log q_t = A + (1-\nu)\log q_{t-1} + \beta_0 \nu \log p_{t-1} + \beta_1 \nu \log y_t + \beta_2 \nu \log w_t + \nu\epsilon_t \tag{9}$$

where $A = \nu \log \beta$, upon linearizing (7) using a simple logarithmic (Napierian) transformation. The price term is lagged one period to reflect the billing procedure of the utility. The consumer does not know until the end of the period what the average price paid was.

This logarithmic flow adjustment model was fitted to the monthly peak system load data previously outlined. The regression results are presented below. The values in parentheses are the standard errors of the estimates.

$$\log q_t = \underset{(1\cdot1505)}{-3\cdot4322} + \underset{(0\cdot0863)}{0\cdot4403}\log q_{t-1} \tag{10}$$

$$\underset{(0\cdot0508)}{-0\cdot2067}\log p_{t-1} + \underset{(0\cdot0138)}{1\cdot8223}\log y_t$$

$$\underset{(0\cdot0009)}{+\ 0\cdot0042}\log w_t \quad R^2 = 0\cdot9978.$$

An estimate of the Durbin–Watson statistic is not reported because of its well known deficiencies with a lagged endogenous variable [6, pp. 298–9].

Consumer response to price changes is significant as is response to income and weather changes. If one gives the price coefficient an elasticity interpretation, one can observe that the price elasticity of −0·207 of the peak system load is much less than the average system elasticity obtained previously of −0·497 (see [12]). It should be noted that price effects could be different at different times of the year, though this is not explicitly included here. If there is a significant space heating or air conditioning load, price responsiveness will vary according to the season of the year. These types of loads, however, are not characteristic of the service area of the utility whose data are used here.

(*c*) *A hybrid model*

Using the prewhitened residuals obtained by the Box–Jenkins technique, a model of the form (3) was fitted to these residuals. In particular, the model

$$a_t = C + \alpha_0\Delta p_{t-1} + \alpha_1\Delta y_t + \alpha_2 w_t + \mu_t \tag{11}$$

where Δ denotes the difference between the current and previous period value of the variable was fit to the same time series. Because in the Box–Jenkins estimation, thirteen of the residuals are lost, the model was actually estimated with data covering the period February 1962 to December of 1969. The variables are as previously defined.

The regression results are presented below. The values in parentheses are the standard errors of the estimates.

$$a_t = \underset{(0\cdot0032)}{-0\cdot0013} \quad \underset{(0\cdot0063)}{-0\cdot0144}\ \Delta p_{t-1} \tag{12}$$

$$\underset{(0\cdot0008)}{+\ 0\cdot0021}\ \Delta y_t + \underset{(0\cdot0003)}{0\cdot0011}\ w_t \qquad R^2 = 0\cdot6705.$$
$$\text{d.w.} = 2\cdot098$$

Price changes, income changes, and weather variations are all significant in explaining the movement of the residuals. The constant term is not significantly different from zero reinforcing the contention above that the residuals have a mean of zero. One should be careful to realize that the coefficient estimates do not have an economic meaning as they did with the econometric model where

the coefficients are elasticities. One can test the hypotheses that the various variables are statistically significant in explaining the movement of the residuals.

IV. FORECASTING WITH THE MODELS

An objective test for determining the superiority of one model over the others can be made by forecasting with the models over a horizon and comparing the forecasts with the realized values.

Recall that the parameters of the model were estimated with monthly data covering the period January 1961 to December 1969. What is done now is to forecast the monthly peak system load over the period January 1970 to December 1973 using the estimated relationships. In the econometric portions, the actual monthly observations on the independent variables are used. It should be emphasized that the econometric model and the hybrid model both have the advantage over the Box–Jenkins model of incorporating actual observations. In practice these values would have to be exogenously forecast which in turn will introduce some additional forecast error. The Box–Jenkins forecasts are made from a single origin.

The forecast results, together with the actual peak system load observations, are presented in Table 2. The problem yet to be resolved is how to define an appropriate measure of forecasting accuracy. While there are many, the one adopted here is the simple absolute deviation from the realized value[7].

The results indicate that the hybrid model does better in 34 of 48 periods than the other models. The econometric model does best in 11 periods and the Box–Jenkins model does best in 3 periods. When compared pairwise, the hybrid model out performs the Box–Jenkins model in 43 of 48 periods and it out performs the econometric model in 37 of 48 periods. The Box–Jenkins model out performs the econometric model in 30 of 48 periods.

The superiority of the hybrid model is well established in this particular instance. Due to an increase in the level of economic activity beginning in mid-1971 that the Box–Jenkins model is not able to reflect and to which the econometric model over-reacts, the hybrid model captures the best of both approaches by adjusting the Box–Jenkins results to reflect this increase.

V. CONCLUDING OBSERVATIONS

It has recently been shown that the Box–Jenkins approach is superior to an econometric approach in forecasting over a relatively short horizon (e.g. [5]).

[7] Forecasting accuracy $= |Z_t - Z_{tr}|$ where $\hat{Z}_t$ = forecast value, and Z_{tr} = realized value.

TABLE 2. FORECAST VALUES (in MW)

Month	Actual	Box–Jenkins	Econometric	Hybrid
January 1970	8628	8653	8461	8604
February 1970	8207	8358	8165	8360
March 1970	8051	8036	8080	8044
April 1970	8276	7841	8062	8084
May 1970	8585	8471	8055	8511
June 1970	9416	9009	8415	9328
July 1970	9697	9473	8520	9502
August 1970	9813	9380	8734	9684
September 1970	9115	8883	8888	9036
October 1970	8524	8354	8475	8533
November 1970	8800	8694	8351	8775
December 1970	9200	9086	8720	9116
January 1971	9000	9034	8767	8996
February 1971	8807	8726	8986	8746
March 1971	8648	8389	8804	8500
April 1971	8390	8186	8796	8347
May 1971	8438	8843	8739	8848
June 1971	9748	9406	8989	9747
July 1971	10271	9890	9490	9990
August 1971	10726	9793	9791	9824
September 1971	10257	9274	10229	9800
October 1971	8769	8722	10054	8744
November 1971	8986	9077	9465	9081
December 1971	9658	9485	9746	9509
January 1972	9546	9432	9899	9476
February 1972	9129	9110	9676	9156
March 1972	9158	8759	9571	8953
April 1972	8927	8546	9944	8613
May 1972	9995	9233	9902	9619
June 1972	11248	9820	10466	9982
July 1972	11655	10326	11090	10553
August 1972	11077	10225	11490	10840
September 1972	10126	9683	11481	9996
October 1972	9413	9106	10734	9442
November 1972	9852	9477	10643	9630
December 1972	10677	9903	10984	10600
January 1973	10245	9847	10966	10189
February 1973	9578	9511	10628	9604
March 1973	9475	9145	10224	9249
April 1973	9514	8923	10188	9228
May 1973	11302	9639	10416	9786
June 1973	12001	10252	11595	11303
July 1973	12089	10781	11694	11305
August 1973	11517	10675	11663	11848
September 1973	9993	10109	11510	10133
October 1973	9539	9506	10386	9549
November 1973	10169	9894	10446	9952
December 1973	9448	10339	10566	10354

The results here support this contention. A combination of the two approaches, however, proves to be clearly superior to either one separately. By taking account of changes in economic and weather related variables in a time series model, improved forecasts are obtained.

REFERENCES

1. Box GEP and Jenkins GM (1970) *Time Series Analysis, Forecasting and Control.* Holden-Day, San Francisco.
2. Bureau of labor statistics (1975) *Employment and Earnings.* US Government Printing Office, Washington.
3. Granger CWJ (1976) Some recent developments in forecasting techniques and strategies, in *Proceedings on Forecasting Methodology For Time-of-Day and Seasonal Electric Utility Loads* (ed. Boyd JW), Electric Power Research Institute, Palo Alto.
4. Houthakker HS and Taylor LD (1970) *Consumer Demand in the United States: Analysis and Projections.* Harvard University Press, Cambridge.
5. Ibrahim IB and Otsuki T (1976) Forecasting GNP components using the method of Box and Jenkins. *Southern Economic J.* **42**(3), 461–470.
6. Koutsoyiannis A (1973) *Theory of Econometrics.* Barnes and Noble, New York.
7. Naylor TH, Seaks TG and Wichern DW (1972) Box–Jenkins methods: an alternative to econometric models. *Int. Stat. Rev.* **40**(2), 123–137.
8. Nelson CR (1974) The first-order moving average process. *J. Econometrics* **2**(2), 121–141.
9. Peschan J and Isaksgen L (1974) *Electric Power Systems Analysis Research.* National Science Foundation, Washington.
10. Smith VK and Cicchetti CJ (1975) Measuring the price elasticity of demand for electricity: the US experience, in *Energy Systems: Forecasting, Planning and Pricing* (Ed. Cicchetti CJ and Foell WK), University of Wisconsin Press, Madison.
11. Uri ND (1976) A new approach to load forecasting in the electrical energy industry. *Cycles* **27**(3), 59–62.
12. Uri ND (1976) A dynamic demand analysis for electrical energy by class of consumer. *Atlantic Ec. J.* **IV**(1), 33–38.

ADDRESS FOR CORRESPONDENCE: *Noel D Uri, Esq., Federal Energy Administration, Office of Coal, Nuclear and Electric Power Analysis, Electric Power Analysis Division, Washington, DC* 20461, *USA.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

Work Force Scheduling with the Search Decision Rule

A DALE FLOWERS

Texas Tech University

STEPHEN E PRESTON

Mostek Corporation, Carrollton, Texas

(*Received November* 1976; *in revised form March* 1977)

1. 5
77

This paper presents the results of a simulation study of the application of the search decision rule to work force scheduling for a manufacturing firm. The problem is of a mixed structure which includes fixed charge components (hiring and training costs), integer variables (work force size in each period), and interacting variables (work force changes, hiring and training costs, and overtime costs). With these elements present, the search decision rule produced lower total costs than the actual company decisions had over the simulated time period. The technique was found to be (1) flexible, (2) easily understood and implemented, and (3) computationally efficient, requiring only a few seconds of computer time per run.

ONE OF the key difficulties encountered in the application of management science techniques do real world operating problems is tha tmany practising managers do not understand the mathematics involved in many of the more sophisticated techniques. This lack of understanding poses a serious threat to actual implementation. One technique for operations scheduling which is relatively easy to understand is the search decision rule (SDR) developed by Taubert [9]. Buffa and Taubert [3, pp. 257–258] have elaborated the advantages and disadvantages of the SDR methodology. One major advantage of the methodology is that it does not require mathematical functions (objectives or constraints) to have restrictive mathematical properties (continuity, differentiability, etc.). Thus, the method is not only more easily understood by the practising manager, but also may be applied to a much wider array of problems than most existing methods. The SDR has the obvious disadvantage that it does not guarantee a mathematically optimum solution. However, Lee and Khumawala [7] compared the SDR to the linear decision rule of Holt *et al.* [4], the management coefficients model of Bowman [2], and the parametric production

planning model of Jones [6], and concluded that for a variety of experimental conditions the SDR clearly outperformed the other models.

The SDR approach is simply the application of the multi-variable search technique developed by Hook and Jeeves [5] to the production scheduling problem. In addition to Taubert's application of this search technique to the scheduling problem, Berry and Bliemel [1] successfully applied it to the problem of finding appropriate smoothing constant values in exponential smoothing forecasting models. In spite of such successes, there have been only a few reported applications of the procedure. In this paper, we simulate the application of the SDR to a department in a job shop manufacturing firm. The problem addressed is one which does not lend itself to standard analytical techniques due to certain problem characteristics.

In our report of this approach, we begin by presenting a description of the firm and the scheduling problem as defined by the managers of the firm. With this information, a mathematical statement of the problem is made followed by an explanation of how the SDR was adapted to fit this problem. Next, the simulation experiments are described and their results presented. Finally, we indicate some possible extensions of this research.

Vol. 5
1977

PROBLEM DESCRIPTION AND FORMULATION

The company involved in the study manufactures a wide variety of tank trailers capable of transporting materials ranging from water to high explosives. Production is primarily on a custom job shop basis although some standard product designs do exist. The total work force of the company varies from 250 to 350 employees depending on demand. For this study, we worked with only one department which had a workforce level ranging from 14 to 23 men. During repeated meetings with company representatives, it was determined that certain cost elements common to many production scheduling problems were negligible or irrelevant for this particular problem. Included in this category were:

(1) layoff costs;
(2) voluntary termination costs;
(3) set up costs;
(4) back order costs;
(5) idle time costs;
(6) inventory associated costs;
(7) subcontracting.

It should be noted that inventory associated costs (especially material costs) were extremely important, but were not significantly affected by the work force scheduling decisions because of the backward due date planning techniques utilized by the shop.

Costs that were important in this analysis included regular time production costs, overtime production costs, and hiring and training costs. Company data were collected for a 65 week period which were used to compute values for each of these costs. It is a company policy to provide a 40 hr work week for each employee. Allowing for unavoidable absences (sickness, etc.), the average hours worked per week per employee was 37·75. Similarly, the average wage earned per hour for regular time production by an employee was $3.86. Because of the company policy mentioned above, the work force level was an integer valued variable. Thus, the total regular time production costs in a given week were computed by multiplying the hours worked times the hourly wage times the work force level.

The overtime production costs were computed by multiplying the overtime hours worked in a particular week by the overtime rate per hour (one and one-half times the regular time rate). Of course, overtime will only be utilized if the scheduled hours to be worked cannot be met with regular time alone.

This brings us to an important characteristic of our problem which we believe to occur fairly frequently in many organizations. In the custom job shop, we could assume perfect forecasts of demand since production was almost exclusively to order. Utilization of the materials requirements planning concept of Orlicky [8] for control purposes also assumes firm schedules. We know of several mass production firms for which this is the case as well as some governmental agencies (certain branches of the census bureau, for example). Thus, it appears that assuming perfect forecasts is reasonable in certain situations. Given this schedule (perfect forecast), the objective in our problem is to find a satisfactory work force level to meet the schedule. This is an intermediate problem between that of aggregate and detailed (shop floor) scheduling, but is structurally more similar to the aggregate problem. In the analysis presented here, it should be remembered that the actual hours worked were taken from the historical data and input to the SDR model as the scheduled hours to be worked.

Because of the perfect forecast situation, we needed only to determine if the work force level times the hours worked per employee met or exceeded the scheduled hours (idle time occurred if it exceeded the scheduled hours, however, this was absorbed by the company). If this number was less than the scheduled hours, the remainder of the hours were computed as overtime. This forced the hours worked to equal the perfect forecasts (schedules). A zero-one variable was utilized in the overtime cost component of the model for accumulating overtime costs.

Due to the resulting problem structure, the work force levels for each period in the planning horizon were the only real decision variables. The regular time production and overtime production costs as well as hiring and training costs were all *residual* variables completely determined by the work force variable. This greatly reduced the problem of solving the model by reducing the number of variables.

The combined hiring and training costs were determined by the company to be \$718.00 per man. These costs are incurred only if the work force in one period has been increased in size from the previous period. Again, a zero-one variable was defined for accumulating hiring and training costs. These costs are high due to the extremely tight labor market in which the company operates. The average unemployment rate for the city in which the plant is located was 4·0% for 1975 and 3·5% for 1976 (source: Texas Employment Commission). Data from these 2 years were used in this study. Generally, then, laid off employees are not available for rehire since other jobs are available. Therefore, complete hiring and training costs are the norm when the company increases the work force size.

The overtime costs and hiring and training costs cause this problem to be similar to the fixed charge problem. An interesting difference is that the zero-one variables utilized in this study are multiplied by expressions which may represent variable amounts for both overtime and hiring and training costs rather than constant amounts. The addition of an integer variable (work force size) further complicates the problem. This problem of mixed structure is stated mathematically in the Appendix.

ADAPTING THE SDR TO THE PROBLEM

Adapting the SDR to the particular problem under study was relatively simple once the problem formulation was complete. The work force level (W_t) variables were the search variables. Since the company uses a 3 month short term planning cycle, the problem was solved for work force variables for 13 weeks.

To utilize the SDR, only a main program and an objective function evaluation subroutine were programmed since the standard search routine is provided in Buffa and Taubert [3]. The main program reads all necessary data inputs and calls the search routine. The search routine then calls the objective function routine as necessary. In the objective function routine, the values of the zero-one variables were determined through a series of logical statements, and then one total cost function was used to accumulate costs in a given time period.

THE SIMULATION EXPERIMENTS

There are several points to be made concerning the simulation experiments. As indicated earlier, we solved for the work force variable for thirteen periods in the future since the company planned on a quarterly basis. In consultation with the managers, we decided that work force sizes should be frozen for 4 weeks into the future and then replanning (solving the model again) should be

performed. This was due to the fact that as new orders came in and production schedules were slipped or accelerated, replanning became necessary. Thus, the first simulation run was for weeks 1–13 of the 65 week period, the second run included weeks 5–17, etc. The solutions to period 1–4 from the first run were used for the SDR decisions while the results from the second run (5–17) determined the SDR decisions for periods 5–8, etc.

In utilizing the search decision rule, some starting values of the decision variables must be supplied for each run. To begin the analysis, the actual company decisions were input to the SDR for the first 13 periods initial solution. After that, the SDR solution from the previous run was used to begin the replanning process for the periods common to both planning horizons. The last overlapping period was used as the solution for all the new periods added to a given run. For example, if the first thirteen work force levels were 15, 15, 15, 15, 15, 14, 14, 14, 16, 16, 16, 16, 16 from the SDR, the first four would be used as firm SDR decisions; the last nine would fill the first nine positions of the beginning solution vector for the next set of computations. The ninth position value was duplicated for position 10–13 for a beginning solution vector. In the example, then, the initial solution vector for run two would be 15, 14, 14, 14, 16, 16, 16, 16, 16, 16, 16, 16, 16.

The beginning work force size for each planning period had to be input to determine if any hiring was necessary in the first period of the planning horizon. For the initial run of the SDR model, the actual work force the company had employed was input for this purpose. Subsequently, the last SDR decision value from the previous run was input as the beginning work force for the succeeding run. In the example above, the size of the work force in period four (15) would be input to the analysis for period 5–17.

The results of the last simulation run for periods 53–65 were used as the SDR decisions for the last thirteen periods since information on scheduled hours for future periods (beyond period 65) were not available for input to this analysis. All these conditions and procedures plus the perfect forecast conditions were deemed to be reasonable by the company managers.

The results of the experiments may now be summarized. The SDR resulted in a total cost of $193,799.23 as opposed to the actual costs of $200,014.09, a difference of $6214.86. These actual costs were taken from company records and are based on differential wage rates for employees whereas the SDR utilized the weighted average of these rates for all periods. To insure that this difference did not affect the comparison of the two sets of decisions, the total cost of the actual company decisions using the same costs as the SDR was computed to be $200,609.88 which results in a slightly more favorable comparison for the SDR.

These cost differences were due primarily to the work force smoothing which the SDR accomplished. Although the SDR utilized perfect forecasts in these analyses, this was deemed a reasonable assumption by the company executives.

Those same executives indicated that they considered the approximately three per cent savings significant since (1) labour costs are a very significant component of variable costs over which some control may be exerted, (2) the company goal is a 10% before tax return (which makes even small percentage decreases in cost important), and (3) the SDR may be applied to other departments to determine if additional savings are possible (recall that the department studied involved 14–23 of a total work force of 250–350).

The algorithm is very efficient since for the fourteen runs used to generate the SDR results an average execution time of 11 sec per run was observed on an IBM 370/145 computer. For all fourteen runs combined, this would result in total computer costs of $13.00 at standard costing rates. The ability to reduce the problem to a search for one set of variable values (work force levels) contributed to this efficiency since dimensionality may create difficulties for the SDR as the size of the problem increases.

SUGGESTIONS FOR FURTHER RESEARCH

The most obvious area for further study in this case is extending the model to include the entire work force. This suggestion was made by company officials. It will involve further complications since labour transfers will have to be considered. Several other costs (termination, etc.) which are not that significant independently, may become so in combination and therefore deserve further study. A final area for research resulting from this study which could have broader implications than the scheduling problem *per se* (this is not to deemphasize the importance and pervasiveness of operations scheduling) is the investigation of the pattern search technique for the solution of fixed charge problems.

Vol. 5
1977

ACKNOWLEDGEMENT

We wish to express our appreciation to the company officials for allowing us access to company data with which to conduct this study.

REFERENCES

1. Berry WL and Bliemel FW (1974) Selecting exponential smoothing constants: an application of pattern search. *Int. J. Prod. Res.* **12**, 483–499.
2. Bowman EH (1963) Consistency and optimality in managerial decision making. *Mgmt. Sci.* **9**, 310–321.
3. Buffa ES and Taubert WH (1972) *Production-Inventory Systems: Planning and Control* (Revised Edition). Irwin, Homewood, IL.

4. HOLT CC, MODIGLIANI F, MUTH JF and SIMON HA (1960) *Planning Production Inventories and Work Force*. Prentice-Hall, Englewood Cliffs, NJ.
5. HOOK R and JEEVES TA (1961) 'Direct search' solution of numerical and statistical problems. *J. Ass. Comput. Mach.* **8**, 212–229.
6. JONES CH (1967) Parametric production planning. *Mgmt. Sci.* **13**, 843–866.
7. LEE WB and KHUMAWALA BM (1974) Simulation testing of aggregate production planning models in an implementation methodology. *Mgmt. Sci.* **20**, 903–911.
8. ORLICKY J (1975) *Materials Requirements Planning*. McGraw-Hill, New York.
9. TAUBERT WH (1968) A search decision rule for the aggregate scheduling problem. *Mgmt. Sci.* **14**, 343–359.

ADDRESS FOR CORRESPONDENCE: *A Dale Flowers Esq, Associate Professor, Texas Tech University, College of Business Administration, Box* 4320, *Lubbock, Texas* 79409, *USA.*

APPENDIX

MATHEMATICAL FORMULATION OF THE MIXED STRUCTURE PROBLEM

$$\text{MIN: } Z = \sum_{t=1}^{N} H_R C_R W_t + \sum_{t=2}^{N} (W_t - W_{t-1}) Z_{1,t} C_H$$

$$+ \sum_{t=1}^{N} (S_t - H_R W_t) Z_{2,t} C_V.$$

$$\text{S.T. } H_R W_t + Z_{2,t} (S_t - H_R W_t) \geq S_t \quad t = 1, \ldots, N$$

$$Z_{1,t} = \begin{cases} 1 \text{ if } W_t > W_{t-1} \\ 0 \text{ otherwise} \end{cases}$$

$$Z_{2,t} = \begin{cases} 1 \text{ if } S_t > H_R W_t \\ 0 \text{ otherwise} \end{cases}$$

W_t integer for all t

where:

N = the number of periods in the planning horizon;
H_R = hours worked per period per employee on regular time;
C_R = cost per hour of regular time production;
W_t = work force size for period t;
$Z_{1,t}$ = a zero-one variable for accumulating hiring and training costs when they occur;
C_H = the hiring and training costs for one employee;
S_t = the scheduled hours (perfect forecasts) to be worked during period t;
$Z_{2,t}$ = a zero-one variable for accumulating overtime costs when they occur;
C_V = cost per hour of overtime production.

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

Screening of Multiattribute Alternatives

RAKESH K SARIN

Woodward-Clyde Consultants, San Francisco

(*Received November* 1976; *in revised form January* 1977)

A procedure for screening multiattribute alternatives is proposed. This procedure reduces the detailed data collection needs and the information burden on the decision maker. It is shown that the inferior decision alternatives can be excluded without specifying probabilities, importance weights, and utilities with precision.

INTRODUCTION

NUMEROUS real life decision situations involve the selection of a decision alternative from a finite number of available alternatives. In this paper we consider an important class of problems in which each alternative is to be judged or evaluated on multiple attributes and the outcome (score) of the alternatives on some or all of the attributes is uncertain.

A common approach to deal with this class of problems is to construct a utility function over the multiple attributes, obtain a probability distribution over the outcomes of the alternatives, and use 'expected utility' as a criterion to rank the alternatives (see Keeney and Raiffa [3] for an excellent discussion of this approach). The complete specification of multiattribute utility function and the elicitation of probabilities, however, requires considerable information from the decision-maker. Moreover, the data collection for specifying the outcome of each alternative may be time consuming and expensive. For example, in a site selection decision situation detailed site specific studies for every site under consideration may be quite expensive.

We propose a screening procedure that reduces considerably the data collection and the effort required of the decision-maker in identifying a preferred decision. This procedure can be implemented conveniently on an on-line interactive computer.

THE PROBLEM STATEMENT

Consider a decision situation in which a decision maker has to choose a preferred alternative from N available alternatives (actions, projects, or

strategies). The outcome of each alternative is evaluated on n attributes (criteria or objectives). This outcome is, however, dependent on the occurrence of some external random events, $E_1, E_2, \ldots, E_m$ only one of which would actually occur. The outcome or score of an alternative k on an attribute i when the random event E_j occurs is denoted by x_{ij}^k.

In order to select a preferred alternative the importance weights for each attribute (w_i), the probabilities of the occurrence of random events when an alternative k is selected (p_j^k) and the conditional utility functions for the attributes (h_i) need to be specified. Once these parameters are specified, the expected utility of each alternative can be computed and the alternative with the highest expected utility is selected; this computation is easier when the attributes are independent [1,2]. The expected utility of an alternative k can be written as

$$\sum_{j=1}^{m} \sum_{i=1}^{n} p_j^k \cdot w_i \cdot h_i(x_{ij}^k) \; \forall \; k = 1 \text{ to } N, \text{ where } \sum_{i=1}^{m} w_i = 1$$

and the conditional utility functions, h_i, are scaled such that the utility of the best score (x_i^*) and the worst score ($x_{i\bullet}$) are 1 and 0 respectively.

Keeney and Raiffa [3] discuss several procedures for obtaining the importance weights and the conditional utility functions. Raiffa [4] and Schlaifer [6] give the methods to elicit subjective probabilities of the decision-maker. We shall assume here that an appropriate procedure has been used to elicit the utility functions and the subjective probabilities.

Vol. 5
1977

THE SCREENING PROCEDURE

The central idea of the proposed screening procedure is to solicit information sequentially until a preferred choice is determined. The sequential strategy involves two phases. In one phase the analyst (or computer) performs analysis to exclude some decision alternatives. In the other phase, the decision-maker supplies some information to the analyst. These phases can be coordinated on an interactive, on-line computer terminal, where the computer does the analysis and queries the decision-maker for certain decisions as and when required. Based on these decisions, further computations are made, and the procedure terminates when a preferred decision is identified.

We shall first describe the procedure in steps of an algorithm for a simple case when the probabilities and the importance weights are exactly specified. An example is provided to illustrate the procedure. The procedure is then generalized to incorporate incomplete knowledge of the importance weights and the probabilities.

Step 1—*initialization*. Obtain w_i for $i = 1$ to n; obtain p_j^k for $j = 1$ to m; and identify $x_{i\bullet}$ and x_i^* for $i = 1$ to n. Essentially, in this step, probabilities are assessed and the worst and the best scores on each attribute are identified.

Step 2—*bounding*. Compute the lower and upper bounds on the utilities of alternative k under event E_j, i.e.

$$\text{Lower bound} = A_j^k = \sum_{i=1}^{n} w_i \cdot f_i(x_{ij}^k),$$

$$\text{Upper bound} = B_j^k = \sum_{i=1}^{n} w_i \cdot g_i(x_{ij}^k),$$

where, $f_i(x_{ij}^k)$ and $g_i(x_{ij}^k)$ are the lower and upper bounds respectively on the utility of the k^{th} alternative's score on the i^{th} attribute under the event E_j. At the start of the procedure, set

$$f_i(x_{ij}^k) = \begin{cases} 1 \text{ if } x_{ij}^k = x_{i*}, \\ 0 \text{ otherwise,} \end{cases} \quad \text{and} \quad gi(x_{ij}^k) = \begin{cases} 0 \text{ if } x_{ij}^k = x_{i*}, \\ 1 \text{ otherwise.} \end{cases}$$

This simply means that the lower bound on the utility of a score is 1 only if it is the best score; otherwise, it is 0. Similarly, the upper bound on the utility of a score is 0 only if it is the worst score; otherwise, it is 1.

Step 3—*testing*. If $Z_{gk}^* \geqq 0$ then alternative g is preferred to alternative k and alternative k can be dropped from further consideration, where

$$Z_{gk}^* = \sum_{j=1}^{m} \left(A^g p_j^g - B_j^k p_j^k \right).$$

The above test checks whether the lower bound on the utility of an alternative is larger than the upper bound on the utility of some other alternative; if it is, the latter is ruled out.

Step 4—*selection*. A score is selected for evaluation by the decision-maker. Several rules for selecting a score are possible; one simple rule is to pick the attribute with the highest weight, w_i, and select a score that has not yet been evaluated. This simple rule performs remarkably well as compared with more elaborate rules (see [5] for experimental comparisons).

Step 5—*evaluation*. The selected score in Step 4 is evaluated by the decision-maker. Based on this information, the bounds in Step 2 are revised and the procedure is repeated until a preferred alternative is identified. Any additional information about the shape of the f_i and g_i functions, e.g. monotonicity, concavity, etc. can be easily incorporated to provide tighter bounds [5]. Similarly, in some situations it is reasonable to assume that a given f_i and g_i is of exponential or some other special form. In these situations as few as one judgement from the decision-maker may be sufficient to completely specify f_i and g_i [3]. This information can also be easily incorporated to provide tighter bounds in Step 2 since $f_i = g_i$.

AN ILLUSTRATIVE EXAMPLE

Suppose there are four available programs for juvenile drug treatment and rehabilitation. The performance of these programs is to be measured on six

attributes. Table 1 lists the criteria and the unit of measurement of these criteria.

TABLE 1. JUVENILE DRUG TREATMENT AND REHABILITATION PROGRAMS

Attributes	Unit of measurement
1. Recidivism	Percentage return to the system in one year of those who completed the program.
2. Increase in maturity	Measured on a 5–point scale.
3. Improvement in family ties	(poor,...,excellent) (5 point)
4. Evaluation of the program by the users	(satisfied, neutral, unsatisfied).
5. Improved relationship with school	(poor,...,excellent) (5 point).
6. Cost of the program	(in '000 dollars).

TABLE 2. PROJECT SCORES TABLE

		Programs[1]			
w_i	Attribute	1	2	3	4
0·3	1	30(0·5), 40(0·5)	50(0·4), 60(0·6)	20	40(0·8), 50(0·2)
0·2	2	2	3(0·5), 5(0·5)	4	2(0·6), 4(0·4)
0·05	3	2	5	1	1
0·15	4	1	3	5	1
0·2	5	3	4(0·5), 5(0·5)	4	2
0·1	6	150	140(0·5) 160(0·5)	130	120

[1] $x_i^k(p)$ indicates probability p of obtaining a score of x_i^k.

In Table 2 the performance of each of the programs on the six attributes is given. For simplicity, in Table 2, the non-numerical scales like poor . . . excellent are assigned numerical values. For example, a score of 1 on the attribute 'Improvement in family ties' corresponds to poor performance, and a score of 5 on this attribute corresponds to excellent performance. The figures in parentheses represent the probability of obtaining the score to the left of the parentheses. For example, Program 1 on attribute 1 has a 0·5 probability of 30% recidivism, and a 0·5 probability of 40%. The outcomes of some programs on some attributes are known with certainty, whereas for others the outcomes are uncertain. Table 2 also gives the weights (w_i's) for each attribute, which are obtained by asking questions from the decision-maker.

The proposed procedure is applied to determine the preferred program. In obtaining lower and upper bounds, the utilities of scores are now weighted by the corresponding probability. For example, in Table 3 the lower and upper bounds of Program 2 on attribute 1 are 0 and 0·4 respectively. To obtain this upper bound, the utilities of 50 and 60 are weighted by the probabilities 0·4 and

Vol. 5
1977

TABLE 3. INITIAL ITERATION

		Programs							
		1		2		3		4	
w_i	Attribute	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
0·3	1	0	1	0	0·4	1	1	0	1
0·2	2	0	0	0·5	1	0	1	0	0·4
0·05	3	0	1	1	1	0	0	0	0
0·15	4	0	0	0	1	1	1	0·5	0
0·2	5	0	1	0·5	1	0	1	0	1
0·1	6	0	1	0	0·5	0	0·5	1	1
		0	0·65	0·25	0·77	0·45	0·90	0·175	0·68

0·6. This gives $g_1{}^2 = 1 \times 4. + 0 \times 6 = 0·4$. (Note that the 60% recidivism rate is the worst score, and hence its upper bound is zero.) In Table 3, lower and upper bounds are given, and the test of comparing the lower weighted total with the upper weighted total does not rule out any program. Now, the utility of a 30% recidivism is evaluated as (say) 0·8.

Table 4 gives the revised lower and upper bounds. In the revision of bounds, monotonicity is exploited. For example, once it is known that the utility of 30% recidivism is 0·8, it is obvious that any score worse than 30% cannot have a utility more than 0·8; hence, the upper bound on the utilities of all recidivism

TABLE 4. FIRST ITERATION

		Programs							
		1		2		3		4	
w_i	Attribute	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
		f_i	g_i	f_i	g_i	f_i	g_i	f_i	g_i
0·3	1	0·4	0·8	0	0·32	1	1	0	0·8
0·2	2	0	0	0·5	1	0	1	0	0·4
0·05	3	0	1	1	1	0	0	0	0
0·15	4	0	0	0	1	1	1	0·5	0
0·2	5	0	1	0·5	1	0	1	0	1
0·1	6	0	1	0	0·5	0	0·5	1	1
		A^1	B^1	A^2	B^2	A^3	B^3	A^4	B^4
		0·12	0·59	0·25	0·746	0·45	0·90	0·175	0·62

rates above 30% is 0·8. Similarly the lower bound on the utilities of all recidivism rates below 30% would be 0·8. Again no program can be ruled out and notice that only the bounds for attribute 1 have changed. Now the utility of

40% recidivism is evaluated as (say) 0·7, and the lower and upper bounds are revised.

Table 5 gives the successive evaluations and the revised bounds for the remaining iterations. After five evaluations (as compared with thirteen required by a complete evaluation scheme), it can be concluded that Program 3 is preferred. Notice also that the data, other than the best and the worst scores, were not needed for the attributes 3, 4, 5, and 6.

TABLE 5. SUCCESSIVE ITERATION

	Program							
	1		2		3		4	
Attribute	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound	Lower bound	Upper bound
	f_i	g_i	f_i	g_i	f_i	g_i	f_i	g_i
1	0·75	0·75	0	0·28	1	1	0·56	0·7
Weighted total	0·225	0·575	0·25	0·734	0·45	0·90	0·343	0·59
Evaluate $h_1(50) = 0·2$ (say).								
1	0·75	0·75	0·08	0·08	1	1	0·6	0·6
Weighted total	0·225	0·575	0·274	0·674	0·45	0·90	0·355	0·56
Evaluate $h_2(3) = 0·2$ (say).								
2	2	0	0·6	0·6	0·2	0·1	0·08	0·4
Weighted total	0·225	0·575	0·294	0·594	0·49	0·90	0·371	0·56
Evaluate $h_2(4) = 0·8$ (say).								
2	0	0	0·6	0·6	0·8	0·8	0·32	0·32
Weighted total	0·225	0·575	0·294	0·594	0·61	0·86	0·419	0·556

CONCLUSIONS

A screening procedure is proposed for the evaluation of multiattribute alternatives. The procedure progresses by sequentially soliciting information from the decision-maker. It was shown that the procedure can be used to screen out alternatives when incomplete information on probabilities, importance weights, and utilities is available. In a practical implementation, it would be desirable to use the procedure to exclude the inferior alternatives with reduced effort and the decision-maker supplies complete evaluations for only a relatively small number of alternatives. Thus, the information burden on the decision-maker is minimized and the detailed data collection is reduced.

The precise specification of the probabilities and the importance weights requires difficult indifference judgments from the decision-maker. For example, to obtain the weight on recidivism in the example above, the decision-maker has to supply his judgment on exactly how much cost he is willing to incur to

reduce the recidivism by a certain amount. In many practical situations it is easier for the decision maker to supply a low value which he is certainly willing to spend and a high value above which he would not spend for reducing the recidivism by a certain amount. In such a situation we would get a range of weights, e.g. the weight on recidivism lies somewhere between 0·2 and 0·4. In the Appendix, we show how the proposed procedure can be modified to exclude some alternatives with simpler information like ordinal ranking and interval estimates on the probabilities and the importance weights. The procedure can also provide a 'forward sensitivity analysis' in that, each time an alternative is excluded, the analyst could query the decision-maker for consistency checks and test under what range of parameters the alternative remains excluded.

We conclude by citing a few of the potential applications of the procedure, i.e.

- Site selection;
- New product decisions;
- Criminal justice project evaluation;
- Research and development project selection;
- Executive recruitment and evaluation;
- Evaluation of projects in education and health systems;
- Selection of corporate plans and strategies.

REFERENCES

1. FISHBURN PC (1965) Independence in utility theory with whole product sets. *Ops Res.* **13**, 28–45.
2. FISHBURN PC (1970) *Utility Theory for Decision Making*, Wiley, New York.
3. KEENEY RL & RAIFFA H (1976) *Decisions with Multiple Objectives*, Wiley, New York.
4. RAIFFA H (1968) *Decision Analysis—Introductory Lectures on Choices Under Uncertainty*. Addision-Wesley, Reading, MA.
5. SARIN RK (1975) Interactive procedures for evaluation of multiattributed alternatives. Working Paper No. 232, Western Management Science Institute, University of California, Los Angeles.
6. SCHLAIFER R (1959) *Probability and Statistics for Business Decisions*. McGraw-Hill, New York.

ADDRESS FOR CORRESPONDENCE: *Rakesh K Sarin Esq, Woodward-Clyde Consultants, Three Embarcadero Center, Suite 700, San Francisco, California 94111, USA*

APPENDIX[2]

Exact estimates of probabilities and ordinal ranking of importance weights

Suppose the following ordinal ranking on the importance weights is obtained:

$$w_1 \geqq w_2 \geqq \ldots \geqq w_n.$$

The proposed procedure can be used with the above information by modifying Step 2 as follows:

$$A_j^k = \text{Min} \sum_{i=1}^{n} w_i f_i(x_{ij}^k) \text{ s.t.} \quad \sum_{i=1}^{n} w_i = 1, w_i \geqq w_{i+1}$$

(1) for $i = 1$ to $n-1$, and

$$B_j^k = \text{Max} \sum_{i=1}^{n} w_i g_i(x_{ij}^k) \text{ s.t.} \quad \sum_{i=1}^{n} w_i = 1, w_i \geqq w_{i+1}$$

for $i = 1$ to $n-1$.

The solution of the above linear programs can be simply obtained, i.e.

$$A^k = \underset{t}{\text{Min}} \frac{1}{t} \sum_{i=1}^{t} f_i(x_{ij}^k), \text{ and}$$

$$B_j^k = \underset{t}{\text{Max}} \frac{1}{t} \sum_{i=1}^{t} g_i(x_{ij}^k).$$

Exact estimates of probabilities and interval estimates of importance weights

Suppose the interval estimates on importance weights, $u_i \leqq w_i \leqq v_i$, $i = 1$ to n, are obtained. The Step 2 is now modified to

$$A_j^k = \text{Min} \sum_{i=1}^{n} w_i f_i(x_{ij}^k) \text{ s.t.} \quad \sum_{i=1}^{n} w_i = 1,$$

(2) $u_i \leqq w_i \leqq v_i$, $i = 1$ to n, and

$$B_j^k = \text{Max} \sum_{i=1}^{n} w_i g_i(x_{ij}^k) \text{ s.t.} \quad \sum_{i=1}^{n} w_i = 1,$$

$u_i \leqq w_i \leqq v_i$, $i = 1$ to n.

The above linear programs can be solved by inspection since these are simple 'knapsack' problems.

Interval estimates of probabilities and exact estimates of importance weights

In this situation the Step 3 is modified as follows:

$$Z^*{}_k = \text{Min}\left[\sum_{j=1}^{m} A_j^q p^q - \sum_{j=1}^{m} B_j^k p_j^k\right]$$

[2] See [5] for a more complete discussion on this Appendix.

Vol. 5
1977

$$\text{s.t.} \sum_{j=1}^{m} p^g = 1, q_j^g \leqq p_j^g \leqq r_j^g \text{ for } j = 1 \text{ to } m,$$

$$\sum_{=1}^{m} p_j^k = 1, q_j^k \leqq p_j^k \leqq r_j^k \text{ for } j = 1 \text{ to } m,$$

$$0 \leqq p^g \leqq 1 \text{ and } 0 \leqq p_j^k \leqq 1, \text{ for } j = 1 \text{ to } m.$$

The above linear program can be simply solved by decomposing it into two linear knapsack problems, one of which is minimized and the other maximized. This test is performed for all (g,k) combinations.

It is easy to see that in the above situation if instead of exact estimates of importance weights ordinal or interval estimates are provided, the Step 2 is simply modified according to (1) and (2), respectively.

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

International Register of Theses and Dissertations

The following are obtainable at the MIT Mciroreproduction Laboratory, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02139, USA:

BECKER BF (1970) An Analysis of National Urbanization Problems/Policies and Existing/Proposed 'New Towns'. SM Thesis, MIT, Cambridge, MA.

BEZOARI UC (1973) The Strategy of Penetration of the US Market by Foreign-Based Multinationals—A Comparison and Contrast of Fiat and Olivetti. SM Thesis, MIT, Cambridge, MA.

BLANKSTEIN CS (1972) The Base Level Development Assistance Project—A Managerial Perspective. SM Thesis, MIT, Cambridge, MA.

BORGMAN LW (1971) The Money Supply and Variable Volatility Portfolios. SM Thesis, MIT, Cambridge, MA.

BORZILLERI R (1971) Fairmicco Inc.—A Case Study of an Industry Ghetto Community Enterprise in Washington, D.C. SM Thesis, MIT, Cambridge, MA.

BRASE WC (1970) Opinion Leadership in Concert Audiences. SM Thesis, MIT, Cambridge, MA.

BROOKS KR (1971) Relative Priorities in the Job–Family Relationship. SM Thesis, MIT, Cambridge, MA.

BYKOV AN (1970) On the Question of the Impact of Computer on the Organizational Structure. SM Thesis, MIT, Cambridge, MA.

BYRNE WJ, JR (1970) Real Estate Investment Trusts—An Overview. SM Thesis, MIT, Cambridge, MA.

CARIGNAN MJ (1973) The Determinants of Investment Advisory Fees. SM Thesis, MIT, Cambridge, MA.

CARMICHAEL GJ III (1973) The Accounting Issue of Income Recognition in the Retail Land Development Industry. SM Thesis, MIT, Cambridge, MA.

CHISHOLM RC (1973) Isolated Colleagues in a Government R&D Organization. SM Thesis, MIT, Cambridge, MA.

CONWAY PD (1972) Portfolio Theory and Transaction Costs. SM Thesis, MIT, Cambridge, MA.

CONWAY WT, JR (1972) An Examination of the Robustness and Stability of the Beta Coefficient as a Measure of Systematic Risk in Managed Portfolios. SM Thesis, MIT, Cambridge, MA.

COOPER JL (1973) The Marketing of Adaptive Housing to the Physically Handicapped. SM Thesis, MIT, Cambridge, MA.

CULLOM CD (1972) Communication Frequency and Colleague Role Performance in an R&D Organization. SM Thesis, MIT, Cambridge, MA.

DUCSIK DW (1973) Shoreline for the Public. SM Thesis, MIT, Cambridge, MA.

FRIEDMAN E (1973) Integration of Japanese and American Management Systems at the Executive Level. SM Thesis, MIT, Cambridge, MA.

GARDNER NE (1973) The Role of Foreign Investment in the Development of the Kenyan Economy Since Independence. SM Thesis, MIT, Cambridge, MA.

GOUDSMIT F (1971) The Impact of Boston Little City Halls on City Management. SM Thesis, MIT, Cambridge, MA.

GROS AJ (1973) A Study of First-Time Blood Donors. SM Thesis, MIT, Cambridge, MA.

GUPTA R (1973) An Alternate Multiphase Approach to Foreign Investments in Lesser Developed Countries. SM Thesis, MIT, Cambridge, MA.

HEWITT DO (1973) Computer Security in an Educational Environment. SM Thesis, MIT, Cambridge, MA.

HIATT DB (1973) The Impact of Charter Services on Scheduled North Atlantic Air Travel. SM Thesis, MIT, Cambridge, MA.

HIRSCH GB (1971) Societal Response to Social Problems. SM Thesis, MIT, Cambridge, MA.

HOEY J (1973) Planning for an Effective Hospital Administrated Emergency Ambulance Service in the City of Boston. SM Thesis, MIT, Cambridge, MA.

HORNUS EH (1971) Methods for the Core for Organizational Studies. SM Thesis, MIT, Cambridge, MA.

HOUGH JD (1973) Jamaica—A Study of the Reasons Behind Foreign Investment and the Use of Environmental Analysis in Decision Making. SM Thesis, MIT, Cambridge, MA.

HUGHES ND (1971) An Investigation of the Response of Brand Shares to Advertising. SM Thesis, MIT, Cambridge, MA.

HURD WC (1972) Computer Simulation of Boston Food Cooperatives. SM Thesis, MIT, Cambridge, MA.

JYOTI P (1973) A Study of Management in Nepal: Theory and Practice. SM Thesis, MIT, Cambridge, MA.

KANODIA H (1973) A Portfolio Approach to the Analysis of a Financial Intermediary—Case Study of the Federal National Mortgage Association. SM Thesis, MIT, Cambridge, MA.

KELLY BJ (1973) The External Analysis of Corporate Strategy—Can It Be Meaningful? SM Thesis, MIT, Cambridge, MA.

KING RM (1973) An Analysis of the Development and Administration of Section 314(d) of the Public Health Service Act, as Amended. SM Thesis, MIT, Cambridge, MA.

LACAYO AO (1973) A Model for Planning the Operations of a Trucking Firm in Nicaragua. SM Thesis, MIT, Cambridge, MA.

LASRY SJ (1973) Manufacturing Information Systems and Requirements Planning—An Analytical Survey. SM Thesis, MIT, Cambridge, MA.

LINDEN RH (1973) The Consequences of Effective Utilization Review of the Massachusetts Hospitals—A Case Study. SM Thesis, MIT, Cambridge, MA.

LITCHFIELD JW (1973) Project Management Game. SM Thesis, MIT, Cambridge, MA.

LYNN JM (1972) A Technique for Scheduling Patients to Specialist Consultations in a Group Practice. SM Thesis, MIT, Cambridge, MA.

MAIETTA VJ (1973) Managerial Potential versus Job Mobility. SM Thesis, Cambridge, MA.

MANTILLA BLF (1973) Multidimensional Scaling and the Data Base—Further Empirical Evidence on Related Issues. SM Thesis, MIT, Cambridge, MA.

MAO RYL (1972) A Study in Corporate Divestiture of Profitable Operations. SM Thesis, Cambridge, MA.

MASON SP (1972) The Evaluation and Planning of a Marketing Opportunity for Briquetted Wood Waste. SM Thesis, MIT, Cambridge, MA.

MCDONALD BA (1972) A Systems Approach to Determining Optimal Levels of Integration for Integrated Circuits. SM Thesis, Cambridge, MA.

MEADOR CL (1972) Long Range Planning with an Interactive Computer Based Decision Support System. SM Thesis, MIT, Cambridge, MA.

MEO PM (1971) Australia: A Study of the Progressively Quicker Trickle-Down of Export Earnings, 1861–1900. SM Thesis, MIT, Cambridge, MA.

MIKULSKY JJ, JR (1973) The Implementation and Operation of Computerized Job Banks in Massachusetts. SM Thesis, MIT, Cambridge, MA.

MONET CR (1973) Interactive Debuggers. SM Thesis, MIT, Cambridge, MA.

MONK JA, JR (1971) A Feasibility Study of a Proposed Admissions Program for the Alfred P Sloan School of Management. SM Thesis, MIT, Cambridge, MA.

MORGAN SW (1973) The Effects of a Management Development Program on Managerial Self-Concept. SM Thesis, MIT, Cambridge, MA.

MOUER CL (1973) The Economics of Liquefied Natural Gas. SM Thesis, MIT, Cambridge, MA.

NISHIYAMA T (1973) Effects of the Measurement Interval and the Measurement Horizon on Beta Coefficients. SM Thesis, MIT, Cambridge, MA.

NOZAWA H (1971) Planning Structure and Process in Commercial Banks. SM Thesis, MIT, Cambridge, MA.

PRICE WC, JR (1971) Factors Influencing Mutual Fund Performance. SM Thesis, MIT, Cambridge, MA.

RANKIN DA (1972) A Model of the Cost of Software Development for the Apollo Spacecraft Computer. SM Thesis, MIT, Cambridge, MA.

ROSENBERG RA (1971) Implementation Problems in Management Information Systems. SM Thesis, MIT, Cambridge, MA.

ROWELL CH (1971) The Theory and Practices of Self-Directed Study—A Case Study of the Undergraduate Studies Program. SM Thesis, MIT, Cambridge, MA.

SCHOEN LP (1972) Investigation of the Pricing and Special Characteristics of Dual-Purpose Mutual Funds. SM Thesis, MIT, Cambridge, MA.

SCHROEDER WW III (1972) Lowell Dynamics—Preliminary Applications of the Theory of Urban Dynamics. SM Thesis, MIT, Cambridge, MA.

SCHWARTZ AH (1972) Capital Budgeting Practices in Industry—Case Study—Raytheon Company. SM Thesis, MIT, Cambridge, MA.

SERVIZIO LA (1971) The Effectiveness of the Brazilian Development Banking System—An Analysis from the Borrowers' Point of View. SM Thesis, MIT, Cambridge, MA.

SHERLING GM (1972) Towards a Formalized Classroom Model. SM Thesis, MIT, Cambridge, MA.

SJOGREN JG (1971) Optimizing Capacity Expansion Models Allowing Use of Non-Proportional Operating Cost Functions. SM Thesis, MIT, Cambridge, MA.

SMITH SG (1971) A Graph Theoretic Approach to Interactive Layout Planning. SM Thesis. MIT, Cambridge, MA.

STEINHOFF E (1971) Short Term Manpower Deployment—An Industrial Dynamics Study, SM Thesis, MIT, Cambridge, MA.

SWEETWOOD HL (1971) Expectational Complementarity, Freshman Seminars, and Student Involvement with MIT. SM Thesis, MIT, Cambridge, MA.

THOMA Y (1973) Labor Mobility in Japan. SM Thesis, MIT, Cambridge, MA.

The following are obtainable at the Library, Cranfield Institute of Technology, Cranfield, Bedford MK43 0AL, UK:

CAMPBELL AJ (1972) The Measurement and Control of Production Efficiency. MSc Thesis, Cranfield Institute of Technology.

CHANDLER BM (1973) An Investigation of Techniques to Simplify Component Location on Complex Printed Circuit Boards. MSc Thesis, Cranfield Institute of Technology.

CHICK KJ (1972) The Design of a Scrap Allowance System for the Helicopter Manufacturing Industry. MSc Thesis, Cranfield Institute of Technology.

DIGENIS G (1972) A Distribution Study for Associated Deliveries Ltd, Reading (joint Thesis with NANDA A). MSc Thesis, Cranfield Institute of Technology.

EDWARDS PD (1971) A Comparison Between Incremental Control and Some Established Machining Methods. MSc Thesis, Cranfield Institute of Technology.

FERGUSON NC (1972) The Development Application and Assessment of a Two Axis Computer Numerical Control System. MSc Thesis, Cranfield Institute of Technology.

GILGEOUS VG (1971) Tool Planning and Control. MSc Thesis, Cranfield Institute of Technology.

GRIFFIN D (1973) The Application of a Three-Phase Squirrel-Cage Induction Motor for a Variable Speed Machine Tool Spindle Drive. MSc Thesis, Cranfield Institute of Technology.

HARPER K (1972) The Control over Maintenance Expenditure in a Vehicle Assembly Plant. MSc Thesis, Cranfield Institute of Technology.

HARRIS R (1972) An Investigation into Wage Payment Schemes. MSc Thesis, Cranfield Institute of Technology.

HAWES R (1973) The Effect of Technological Change on Management/Employee Relations. MSc Thesis, Cranfield Institute of Technology.

HIBBERD NJC (1971) An Investigation into the Organisational Problems with Component Coding. MSc Thesis, Cranfield Institute of Technology.

HOLLAND R (1972) Heart Rate Variability—Perceptual Load. MSc Thesis, Cranfield Institute of Technology.

HOWELL CA (1971) An Investigation into the Problems of Introducing Group Technology into a Small Batch Manufacturing Company. MSc Thesis, Cranfield Institute of Technology.

HUNDAL RS (1971) Production Planning and Scheduling Systems for Batch Manufacture to Customers Orders Applied to a Small Firm. MSc Thesis, Cranfield Institute of Technology.

JEFFERIES RS (1972) An Investigation into a Sub-Contract Manufacturing Company. MSc Thesis, Cranfield Institute of Technology.

JOORABCHIAN A (1972) Investigations into the Problem of Sand for an Iranian Foundary. MSc Thesis, Cranfield Institute of Technology.

KENNEDY L (1970) An Investigation into the Use of the Edge-Tone Phenomena in Fluidic Devices. MSc. Thesis, Cranfield Institute of Technology.

KITE DJ (1970) Investigation into the Effect of Tool Geometry on Finish Machining. MSc Thesis, Cranfield Institute of Technology.

LAU RSS (1971) Investigation into the Application of Production Planning and Control to the Maintenance of Aircraft Having Limited and Variable Flying Hours. MSc Thesis, Cranfield Institute of Technology.

LEEK CJ (1972) The Design and Development of a Pneumohydraulic Positioning System. MSc Thesis, Cranfield Institute of Technology.

LEFLEY BN (1970) An Investigation into Tool Tip Temperature Measurement and Tool Wear on a Throwaway Tip Lathe Tool. MSc Thesis, Cranfield Institute of Technology.

LOVE WJ (1971) Three Orthogonal Laser Reference Planes for Alignment Testing with Applications to Machine Tools. MSc Thesis, Cranfield Institute of Technology.

MCCRACKEN SG (1972) Proposal for Total Quality Control in a Large Aircraft Manufacturing Company. MSc Thesis, Cranfield Institute of Technology.

MCGARRELL PJ (1973) An Examination of Information Flow in an Industrial Computer System. MSc Thesis, Cranfield Institute of Technology.

PAYNE DJ (1972) Computer Monitoring of an Automatic Machining Line. MSc Thesis, Cranfield Institute of Technology.

RENDLE J (1973) A Diagnostic Approach to Industrial Engineering—Development of Financial Criteria for Group Technology Feasibility. MSc Thesis, Cranfield Institute of Technology.

RICHARDS JB (1972) An Investigation into the Factors Affecting Operator Replacement. MSc Thesis, Cranfield Institute of Technology.

RODWELL JAW (1973) The Influence of the Dynamic Stiffness of a Planing Machine upon the Workpiece Surface Texture. MSc Thesis, Cranfield Institute of Technology.

SAGHATOLESLAMI M (1972) Developments in the Trim Loss Problem. MSc Thesis, Cranfield Institute of Technology.

SEWARD KJ (1973) An Investigation into Methods of Crack Detection Using Optical Holography. MSc Thesis, Cranfield Institute of Technology.

SHERBINI AAM (1973) Development and Testing of on-Line System for a Computer Controlled Cam Grinding Machine. MSc Thesis, Cranfield Institute of Technology.

SIVANESAN S (1973) An Investigation into Manpower Planning at Company Level. MSc Thesis, Cranfield Institute of Technology.

SLATER M (1973) A Manufacturing System for Precision Rotary Components—An Application of Group Technology. MSc Thesis, Cranfield Institute of Technology.

SPAVEN NF (1973) Group Technology: Its Application and Control Within a Batch Manufacturing Environment. MSc Thesis, Cranfield Institute of Technology.

STANGER AG (1972) Production Planning and Inventory Control in a Small Company. MSc Thesis, Cranfield Institute of Technology.

STERICKER SW (1972) Integration of Hand Tool Design Parameters for Human Performance. MSc Thesis, Cranfield Institute of Technology.

STOTT D (1971) The Application of a Small Computer to the Direct Numerical Control of a Machine Tool. MSc Thesis, Cranfield Institute of Technology.

SWARUP JLS (1973) An Examination of a System for Computerized Data Acquisition in Industry. MSc Thesis, Cranfield Institute of Technology.

TANCOCK DJ (1972) An Investigation into the Future Investment Policy Within a Small Manufacturing Company. MSc Thesis, Cranfield Institute of Technology.

THOMAS PH (1972) A Study of Garment Manufacture (Man-Made Fabric Welding Processes), MSc Thesis, Cranfield Institute of Technology.

THOMPSON PR (1973) An Investigation of a Method of Maintaining Constant Gap in a Single Pad Air Thrust Bearing. MSc Thesis, Cranfield Institute of Technology.

THOROGOOD BP (1972) An Investigation into Human Limitations on Combine Harvester Performance. MSc Thesis, Cranfield Institute of Technology.

TUNLEY J (1973) The Development of an Adaptive Control System for a Lathe Using a Small on-Line Digital Computer. MSc Thesis, Cranfield Institute of Technology.

WALKER LJ (1972) The Effects of Increased Press Output on the Post-Press Operations at an Aluminium Extrusion Plant. MSc Thesis, Cranfield Institute of Technology.

Information for inclusion in this Section, presented in the appropriate form, should be sent to the Editor of Omega.

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 4, 1977. Pergamon Press. Printed in Great Britain

ANNOUNCEMENT

A ONE-DAY CONFERENCE

ON

THEORY AND PRACTICE IN MANAGEMENT SCIENCE

(in conjunction with OMEGA, International Journal of Management Science),
at Imperial College, London

9 September 1977

The Conference will draw together representatives of senior and departmental management, and operational research/management science. The purpose of this conference is to identify factors inhibiting the full use and implementation of OR methods as seen by managers and OR scientists.

Speakers will include:

Mr DAVID BISHOP	Deputy General Manager, Ransome Hoffman Pollard Ltd, Bearings Division
Mr KEN BOWEN	Head of Research, Defence Operational Analysis Establishment
Professor SL COOK	University of Aston in Birmingham
Mr PG FITZGERALD	Serck Limited
Sir FRANK MCFADZEAN	Chairman, British Airways
Mr RONALD UTIGER	Managing Director, British Aluminium Co.
Mr VICTOR WEST	General Manager, Pendle Insurance Co.
Conference Fee:	£20 including lunch and refreshments

For full details and application forms apply to:

Edith Simmons
Editorial Assistant
OMEGA
Department of Management Science
Imperial College
Exhibition Road
London SW7 2BX

BLANK
PAGE

Vol. 5
1977

OMEGA

The International Journal of Management Science
Volume 5, Number 5, 1977

Contents

SAMUEL EILON	499	Editorial—Structural Determinism
	505	Feedback
UMBERTO COLOMBO	511	Strategies for Europe. Proposals for Science and Technology Policies. Industrial Innovation in Europe
HSD COLE	529	Accuracy in the Long Run—Where Are We Now?
RB FLAVELL	543	Divisionalization and Transfer Pricing: A Review
JC HIGGINS, R FINN	557	The Chief Executive and his Information System
JR MACE	567	Criteria for the Selection of Reporting Method
HG JONES	583	Investment and the Swedish Economy
DONALD V MATHUSZ	593	The Value of Information Concept Applied to Data Systems
JC HULL	605	Memoranda
JC HERSHAUER, JW GOWENS	609	
JM GLEASON	611	
DAVID A LOPEZ, WILLIAM J VAN SLYKE	614	
BERNARD W TAYLOR, III, ARTHUR J KEOWN	618	
	621	International Register of Theses and Dissertations

Special Notice

To avoid simultaneous, multiple submissions to journals having similar interests, the Editors of *Management Science*, the *Operational Research Quarterly*, *Operations Research*, *Omega* and *Naval Research Logistics Quarterly* have developed a system of exchange submission lists.

Because refereeing capability is a scarce resource, we cannot tolerate the duplication of effort that follows from multiple submissions. We hope that the elimination of multiple submissions will make the refereeing process more efficient, and thereby decrease the time required for refereeing. This is a major objective of all five publications. Other journals with similar interests may soon be included in this information interchange.

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci, Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

EDITORIAL

Structural Determinism

IT HAD to come. Ever since Joan Woodward published her thesis on the relationship between technology and organization structure [6] there have been many reactions to her work, ranging from glowing endorsements of what some people considered were fundamental and novel contributions to the field of organization behaviour to outright scepticism and refutations. I have to confess that in spite of the immense field work and the very time consuming data collection exercises that she and her team were engaged on, it had been evident to me for some time that her theory would not stand up to a searching examination. It was too simple, too static, too pat, and so, cracks were soon to appear; and now a critique by Lex Donaldson [1] has so damaged its very foundations, that it is doubtful whether the theory in anything resembling its original form can possibly survive.

It must be said that there was something naively attractive about the proposition "that the organizational structure and processes of a manufacturing firm are causally related to its technology", and this was claimed to be "a comparatively new notion, for the ideas about management structure and behaviour developed in the first half of this century were based on a monolithic concept of organization and on the belief that there is one best and all-embracing way to run any manufacturing business" [7, p. ix]. Not so, said Woodward; the lumping together of all manufacturing enterprises for this purpose is unviable, but when manufacturing industry is classified in categories according to the technology employed (specifically, three categories were originally identified, first of firms "producing units or small batches mainly to customers' orders, secondly, large batch and mass production firms, and thirdly, process firms concerned with the continuous production of gases, liquids and crystalline substances" [7, p. ix]), then specific organizational characteristics in the firms concerned can be related to the technology involved.

Woodward was, of course, right to reject the notion that all manufacturing industry can and should be organized in the same way, but she was wrong in thinking that, once manufacturing firms were segregated into three or more categories according to their technology, distinct organizational patterns would emerge. Alas, this simple proposition was not fully supported even by results

from her own empirical work, and she admitted that her "studies had shown that in relating technology to organizational behaviour there are two clear-cut extremes in the technical scale" (i.e. unit or small batch production on the one hand and process or continuous flow on the other) "and a large and difficult centre area in which are found medium to large batch production firms and firms with the component assembly type of production system. Structure and behaviour appear to be more consistent and predictable at the extremes than in the centre" [7, p. x]. Thus, even if we were to accept her conclusions as valid for the extremes, then by her own admission the majority of manufacturing industry remains outside the scope of the technology theory.

But instead of drawing the conclusion that the theory was fundamentally suspect, she postulated that "One possible explanation of the apparent lack of consistency was that the classification of technology (used in the earlier study) was inadequate" and concluded that "one of the tasks of the new research team, therefore, was to try to find a more sophisticated way of identifying technical differences between firms and then to relate these to differences observed in organizational behaviour" [7, p. x]. In other words, she persisted in holding on to her original proposition that organization structure and technology are related, and as the first attempt at a technology classification had not been entirely successful, she inferred that a more detailed classification was bound to lead to success. It was perhaps characteristic of her work that it was the difficulty in categorization of technology, not the validity of the theory itself, that became the focal point of the research effort. But five years later that goal remained as elusive as ever, and she admitted in the conclusion of her last book: "nor is it claimed that there have been any startling breakthroughs in the way of new middle range theories. It is felt, however, that more insights into the relationships between technology and organizational structure have resulted from the study and categorization of control processes, as has the realization that not only the nature of the task, but the way that it is planned and controlled, is an important determinant of organizational behaviour" [7, p. 242], a proposition that had in fact been made to her several years earlier [2].

Vol. 5
1977

But let us start at the beginning: why is it important to debate whether Woodward's theory is valid or not? Is this a matter of some practical significance or is it merely an obscure academic discussion? The answer is, of course, that the theory—if it is correct—lies at the heart of the problems of organizational design and organizational choice. If there is one way in which given firms can or should be organized and if the only or major determinant is the technology with which the firms are associated, then we have a clear recipe to guide us both in those cases where new firms are set up and in those where, for some reason, the firms deviate from the 'optimum' or 'expected' structure.

The question, therefore, is whether the theory is valid. This school of thought, which belongs to the view of *structural determinism* [see 3, Chapter 10], ignores the personalities of the top executives in the managerial system and the possible

effects that they might have on the development and exercise of control procedures; it argues instead that the organization structure of a given firm is the result of a process of adaptation in response to pressures exerted by the environment. Now, this environment consists of several parameters, of which technology is one (others include the market, the trade unions, the economic climate, the industry in which the firm operates, etc.). To say that the organization structure of a firm is technology dependent is merely to assert that technology is always the dominant parameter and that other parameters either play no part at all or are insignificant in their ability to influence the structure. It is an appealing hypothesis, engagingly simple in its concept in that it makes the obvious proposition that managerial control is task orientated, and since manufacturing tasks are inextricably interwoven with the technology used to carry them out, the inevitable effect of technology on organization becomes self-evident.

So far so good; but the technology theory went further in asserting not only that technology does have such an effect but that it is the major (even the sole) determinant. It is easy to visualize why empiricists found such a simple theory so attractive: all they needed was to have a reasonable classification of technology, based on certain technical parameters, and a given characterization of organization structure, based on such measures as the average span of control of first line supervisors and the number of hierarchical levels, and they could then proceed to test whether a relationship between the two holds.

I shall not dwell on the real problems of producing such classifications. Apart from the fact that inferences about a firm as a whole may be questionable when they are based on observations at a micro level (there is a danger, for example, in concluding that a firm belongs to the mass production type just because its final assembly lines are so operated; many firms characterized as mass producers have feeder departments engaged in batch production), there is the problem that both sets of classifications may depend on common attributes by the very way in which they are defined at the outset, so that their interdependence may be a reflection of their definition rather than of empirical results (see [3, p. 198]). Questions about the methodology employed in such empirical work were therefore bound to arise, particularly among investigators attempting to replicate Woodward's results.

Perhaps the first overt challenge to the technology theory (discounting the dissenting comments that I had made to Joan Woodward in many private communications) came from Pugh and his colleagues [4, 5] on the question of the effect of size on organization structure. As a result of her empirical work, Woodward concluded that in the hundred firms in her sample "organizational characteristics were not directly related to size" [7, p. ix], whereas Pugh's results suggested that size was the most predominant variable. Such divergent views clearly suggest that, apart from questions of sample size and considerations of statistical significance, there could be differences in definitions of what is meant by organization structure. For example, if the major measure of structure

is taken as the span of first line supervisors or the vertical span (number of hierarchical levels), then it is theoretically possible for either to remain constant when the size of the organization varies, but it is inconceivable for both to remain unaffected by the dimensions of the pyramid. And since the size of the pyramidal hierarchy is a function of the size of the firm, it is futile to debate whether structure and size are interdependent, if both are defined by common measures. This is why I argued that "what is required is a definition of various types of structure, based on a list of given attributes, and similarly a parallel definition of types of technology, based on a *separate* set of attributes, *prior* to the testing of the hypothesis that a correlation between the two exists" [3, p. 200].

It is not surprising, therefore, that this level of rigour alone should have aroused criticisms of the technology theory, with other investigators failing to replicate Woodward's results. But there is a much more fundamental issue at stake, namely the fact that any theory embedded in structural determinism involves the notion of inevitability in the way that an organization develops, suggesting that a particular technology must result in the organization structure being of a particular type. Leaving aside the batch production area, for which Woodward's findings did not support this conclusion, the implications of such a theory are that irrespective of the product, irrespective of the market place or competition, and irrespective of the personalities of the top managers in the system, the outcome in terms of organization characteristics must be the same, if a given type of production is employed. Such a conclusion simply offends the intuitive belief of many of us in the existence of choice in organization design and the validity of the notion that firms can be restructured as a result of conscious managerial decisions.

Vol. 5
1977

Woodward went, in fact, further than the notion of structural determinism, in that she postulated as a result of her findings that the fit between technology and structure is related to organizational success, or to put it explicitly Woodward argued that while firms in her sample showed that "organizational characteristics were not directly related to . . . the degree of business success they enjoyed" [7, p. ix], she nevertheless concluded that "organizational structure took an optimum form which varied according to the production technology" [1, p. 263]. It would seem, therefore, that the technology theory culminated in two propositions, the first linking structure and technology as an inevitable outcome (except in the large batch production area) and the second linking optimal structure with success. It is difficult to see how the two can be reconciled, unless the inevitably resultant structure is associated with a fairly wide range within which a structural choice is thought to be possible, and this is precisely what Woodward suggested: "It was found that the figures relating to the organizational characteristics of the successful firms in each production category tended to cluster round the medians for the category as a whole, while the figures of the firms classified as 'below average' in success, were found at the extreme of the

range" [6, p. 65]. The weakness of such a compromise is sufficiently apparent to make it an improbable working hypothesis, since the ranges for the various technology groups would then have to be so wide as to result in substantial overlaps [6, Table IV, p. 69; see also 1, p. 264], but the wider the ranges and the more they overlap, the less tenable is the proposed relationship between technology and structure without the support of large samples and statistical tests.

Even if the logical dichotomy between the two propositions of structural determinism and relation to success were somehow to be reconciled, one would have to explain how it is that a particular organization structure allied to a given technology could ensure success without any reference to the market place. To put it crudely, two firms employing the same technology and organized along similar lines need not be equally successful; the ability to make profit (as an example of one measure of success) may well depend on the level of competition and on the pricing policies adopted by the two firms, and it would be far-fetched to argue that their pricing policies would necessarily become identical because their technologies happen to be the same.

It is for all these reasons that the technology theory could not last. What Donaldson has done in his critique is to review the literature in this field mainly since 1965, when Woodward's book was published [6], and specifically to show how four major attempts to replicate her findings have failed. "What are the implications for the management practitioner? Woodward indicated that certain organizational structures are appropriate for certain technologies and, by implication, that performance would suffer if the inappropriate structures were adopted" [1, p. 272]. These conclusions are unwarranted and Donaldson ends his critical review by saying: "Of the bivariate relationship" (structure vs technology and structure vs success) "which Woodward found, none have been confirmed by the four major studies. Whilst further research is undoubtedly warranted it cannot be certain that such research will identify additional variables, or methodological factors, which will explain away the discrepancies between the findings of Woodward and those of the subsequent studies. The plainest way of interpreting the current evidence is that it disconfirms aspects of the original Woodward thesis" [1, p. 273].

SAMUEL EILON
Chief Editor

REFERENCES

1. DONALDSON L (1976) Woodward, technology, organizational structure and performance—a critique of the universal generalization. *J. Mgmt Studies* **13**, 255–273.
2. EILON S (1962) Problems in studying management control. *Int. J. Prod. Res.* **1**, 13-20.
3. EILON S (1971) *Management Control*. Macmillan, London.
4. PUGH DS, HICKSON JD, HININGS CR and TURNER C (1969) Dimensions of organization structure. *Admin. Sci. Q.* **13**, 65–105.

5. PUGH DS, HICKSON JD, HININGS CR and TURNER C (1969) The contest of organization structures. *Admin. Sci. Q.* **14**, 91–114.
6. WOODWARD J (1965) *Industrial Organization—Theory and Practice*. Oxford University Press.
7. WOODWARD J (Ed.) (1970) *Industrial Organization—Behaviour and Control*. Oxford University Press, London.

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

Feedback

On Structural Mapping for Policy Selection

THE SUGGESTION that the result of structural mapping [1] can indicate policy preference on a basis of policy indifference is intrinsically implausible and apparently false. The suggestion is specifically that the mapping will put most and least preferred policies to opposite extremes of the graph of which vertices are policies and edges link policies of similar attractiveness, where similarity is not a transitive relation. Consider any finite planar graph embedded as points and non-intersecting line segments in a plane and having two easily distinguished extreme vertices, which Professor Rivett's suggestion would lead one to expect would be the extremes of preference. Orient the graph so that these two vertices are one above the other. I propose that the indifference links of the graph are as compatible with a decline in preference from left to right, where the preference along a single edge is sufficiently slight (zero for a vertical edge) that they are regarded as indifferent, as with a decline from top to bottom.

The restriction to a strictly planar graph is easily seen not to be essential; it just makes the situation easier to state.

MR THOMAS' comment on my paper is interesting but on reflection it can be seen, I believe, to be misguided. He cites an example where the preference axis can be the second component and not the principal component. Clearly what he suggests is possible but is very unlikely to occur. For his requirement is that if we have (say) five policies, A, B, C, D and E which are (roughly) similar in attractiveness then our linking will be of the form (A,B), (B,C), (C,D) and (D,E) giving

$$A\text{–}B\text{–}C\text{–}D\text{–}E.$$

Any assignment of other links, such as A–C etc. will have the effect of reducing the length of this chain and making it 'fatter'. It has been shown in practice, using MDSCAL, that the more links there are between these five, for example, the more their configuration tends to a small grouping.

The second requirement for Mr. Thomas' rebuttal to work is that there is a second (small) axis which corresponds to the actual choice; for example $C \sim F$ and $C \sim G$, giving a configuration

$$\begin{array}{c} F \\ | \\ A\text{–}B\text{–}C\text{–}D\text{–}E. \\ | \\ G \end{array}$$

This is all clearly *possible*. It requires however that, in this case, of the seven policies five should be of roughly equal merit, with a sixth and seventh rather better and rather worse (but still) linked to one of the five. It also requires that the five be linked in a very special way, namely that although by hypothesis all five are of equal merit, the links perceived by the decision maker are A–B–C–D–E. If p is the probability of a link between any two of (A, B, C, D, E) and $q = 1 - p$, then the probability of such a chain is $60\ p^4q^6$. Since p must be assumed fairly large, the probability of such an occurrence

REFERENCE

1. RIVETT BHP (1977) Multidimension Scaling for Multi Objective Policies. *Omega* **5**(4), 367–379.

RSD THOMAS

Department of Computer Science
The University of Manitoba
Winnipeg
Canada R3T 2N2

as that envisioned by Mr. Thomas must be rather small.

As in all problems where we are dealing with chance we can never exclude catastrophe what might possibly happen will (almost) certainly occur if we wait long enough!

BHP RIVETT

Professor of Operational Research
The University of Sussex
Operational Research
Mantell Building
Falmer
Brighton BN1 9RF
UK

Technicians, Advisers—and All-Rounders in Operational Research?

EILON'S recent editorial [6], 'Technician or Adviser?', has raised again the extremely important question of the proper scope of Operational Research (OR). In making this contribution, to what we hope will be a major debate on the issue, we should stress that our own views are fairly close to those stated by Eilon. We believe that OR should be involved in the important executive problems. For those who disagree we should also state here, at the outset, that what follows will take the reader through paths rarely frequented by the OR community. . . .

Any relatively dispassionate analysis [5] would indicate that decision-makers in modern organisations are likely to encounter a variety of problems, which range in complexity from the well-structured to the very ill-defined. Naturally, the well-structured problems, for which standard techniques are appropriate, tend to be tactical (of limited consequence) whereas the ill-defined, more complex problems tend to be strategic (and therefore of more importance for the organisation).

Quite clearly, at least some of the original OR community considered that it should be involved with the full range of decision problems [1]. If, as we believe, Eilon's description of a single OR group is a fair representation of most practice in the UK today, the question, we think, is not why Eilon made his statement, but why it was necessary for him to make it. Why is it that we have become almost exclusively tactical-problem technicians?

In attempting to answer this question, it is crucial to keep in mind that this restriction, of a community supposedly devoted to the use of scientific method to assist decision-making, has occurred despite the pressing need for such help in many complex and important issues (see, for example, [3]). Although we have heard a number of possible explanations, most have been at an organisational level. A popular example seems to be the simple matter of survival. In order to establish and develop an OR group in an organisation is it necessary to show success. Thus well-structured problems with tangible solutions have to have overwhelming priority. This limited focus of explanation is perhaps to be expected [4]. However, we would like to suggest a factor of a different order and consider the likely consequences of such organisational processes on the composition of the OR community itself.

It is our belief that Blackett's idea of OR [1], as a general scientific approach to analysis for any type of decision-problem, has been essentially dead for twenty years. Since the 1950's the 'new OR man' has been, above all else, the numerate technician who [6]

> ". . . is given a problem and sets about solving it . . . accepts only limited responsibility for the definition of the problems and for their stated objectives, but total responsibility for the execution of his projects, once they are clearly defined . . ."

Moreover, this is the dominant impression to be derived by an outsider examining our methods, journals, conferences and education courses [4]. *We suggest that this state of affairs influenced the people who became OR workers, that what OR did (inadvertently?) defined what it could, and should do, in the minds of potential recruits.*

This scarcely controversial suggestion is important because of the nature of the recruits who were available. Hudson [7, 8] has studied the interests and career choices

Vol. 5
1977

of high-ability British schoolchildren since the late 1950's. What is pertinent here is his differentiation between convergers and divergers [7, p. 42]:

". . . Most arts specialists, weak at the IQ tests, were much better at the open ended ones; most scientists were the reverse. Arts specialists are on the whole divergers, physical scientists convergers. Between three and four divergers go into the arts subjects like history, English literature and modern languages for every one that goes into physical science. And, vice versa, between three and four convergers do mathematics, physics and chemistry for every one that goes into the arts. As far as one can tell from the samples available, classics belong with physical science, while biology, geography, economics and general arts courses attract convergers and divergers in roughly equal proportions."

Convergers are good at, and attracted to, well-defined closed problems. Divergers, on the other hand, are good at, and attracted to open-ended problems (that is, to more complex issues for which there is no single right answer). It is important to understand that these are not trivial differences. They are deep-seated characteristics related to the acceptability of authority [8, p. 5]:

"Both at school and at university this confrontation with intellectual authority is especially acute . . . This massive, largely unavoidable insistence on authoritative knowledge faces the student with an unenviable choice. . . .

The gist of my argument is that convergers and divergers will differ in their reaction to this dilemma. The first will tend to plump for those routes through the academic system—mathematics, physical science, classics—in which the weight of accepted authority is greatest; the second, those in which the pressure is least. What is alluring to the converger (and repugnant to the diverger) about the exact disciplines is their exactitude. They are systems of thought from which both muddle and emotion have largely been removed. . . .

In their choice of a life's work, convergers are drawn towards careers which hold the promise of crystalline exactitude and the ultimate exclusion of doubt. While for divergers, it is not only the emotional connotations of the arts that render them so attractive, but also their imprecision. . ."

In summary, schoolchildren manifest deep-seated differences in interests which are strongly linked to career choices. Convergers move toward mathematics and physical science, divergers do the opposite. All-rounders who manifest mixed abilities (that is, people who are able to tackle closed and open-ended problems well) divide more evenly between subjects.

What then, could we expect a population manifesting such characteristics to have made of an OR community concentrating on numerate analysis of well-structured problems—indeed apparently concerned with little else. Surely, convergers would have been attracted preferentially. Moreover, the process would have built up over the years. The more convergers that were preferentially attracted, the more we would have concentrated on well-structured problems, and the less we would have attracted all-rounders with their divergent abilities. Unfortunately for us this was all taking place (if we are correct) during the pronounced swing away from hard science in the UK school population. This swing, we presume, must have greatly reduced the proportion of all-rounders available as potential recruits to OR in the first place.

Thus it seems probable to us that, in a central respect, we have become almost the opposite of our professed interdisciplinary nature. A positive feedback process has produced a population markedly skewed towards convergers instead of the ideal mixed population with all-rounders predominating. The answer to the original question is therefore that we are technicians because that is what most of us want to be.

The implications for those of us wishing OR to take up its broader role (which we, at least, believe that it must do to survive) are clear and serious. We are asking a convergent population to attempt divergent problems for which they are not temperamentally suited. Above all else what differentiates the OR adviser from the technician, and what is needed in complex strategic problems, is a dislike for accepting what is given. In Eilon's apt words [6]

". . . the adviser, on the other hand, is not interested in a particular problem, but in how it has come about, whether valid questions are being asked, whether any prescribed constraints should be challenged; above all, the adviser is the man who helps to formulate the problem in the first instance, in defining objectives and in

assessing wide ranging implications . . ."
It should not surprise us that such a role does not come easily to a man whose original interest was in "the promise of crystalline exactitude."

Although Hudson is always careful to stress that convergers can, to some extent, become more divergent (if told clearly enough what is required) we do tend to get set in our ways. Changing the present population must be regarded as a difficult, slow and potentially painful process.

However, there is a rather faster road to righting the present imbalance in the OR community—to change our OR education courses. Our own recent experience has been of growing complaints from practical OR men about inappropriate university education courses; and the grotesquely excessive emphasis on mathematical techniques in such courses has been a major factor in the process we have described above. We should get back to teaching about scientific method in general and the context in which we find our own problems in particular. (How can we justify sending students out into the real world with numerous courses in mathematical techniques and almost total ignorance of modern work on politics and power in organisations?) As Bondi [2] has stated (in arguing for more science graduates in executive positions):

> "If one had viewed science as an education of the mind rather than as professional training, naturally students emerging with a first degree would have known less science. But they would have known more about the kingdom of scientific ideas, about how to argue about the evolution of human thinking and about how scientific debate and the interaction of experiment and theory stimulate creativity.
>
> Further, such an orientation of science courses might also produce better scientists, first because such teaching would attract many of the ablest young people now repelled by the supposedly inhuman face of science, and second because in science itself an ability to argue, debate, and understand the other person's ideas, however imperfectly expressed, is as important as sheer scientific knowledge."

To us, the need for such scientists is even more essential in OR, but perhaps we should leave the last word with Blackett [1, p. 201]:

> ". . . In my view the element of relative novelty (in OR) lies not so much in the material to which the scientific method is applied as in the level at which the work is done, in the comparative freedom of the investigators to seek out their own problems, and in the direct relation of the work to the possibilities of executive action. . . ."

and finally [1, p. 203]

> ". . . operational research is scientific, and training in some scientific discipline may be regarded as essential, although it need not necessarily be in the exact sciences. The most important qualification is the ability to take a broad view of the problem, so that the most important factors will not be missed. . . ."

REFERENCES

1. Blackett PMS (1962) *Studies of War*, Part II, *Operational Research*. Oliver & Boyd, Edinburgh.
2. Bondi H (1977) Science as an education. *Nature* **265**, 286.
3. Cook SL (1973) Operational research, social well-being and the zero growth concept. *Omega* **1** (6), 647–667.
4. Dando MR and Eden C (1977) Reflections on operational research: A report from the Euro II congress. *Omega* **5**(3), 255–270.
5. Dando MR, Sharp RG and Defrenne A (1977) Could OR be a science? *Omega* **5** (1), 89–92.
6. Eilon S (1977) Technician or adviser? *Omega* **5** (1), 1–6.
7. Hudson L (1966) *Contrary Imaginations*. Methuen, London.
8. Hudson L (1968) *Frames of Reference*. Methuen, London.

MR Dando
RG Sharp

The University of Sussex
Department of Operational Research
Mantell Building
Falmer
Brighton BN1 9RF
UK

Vol. 5
1977

Omega, Vol. 5, No. 5

The Quality of Working Life—A Comment

IN HIS editorial on the quality of working life Professor Eilon reviews the two volumes on this topic edited by Davis and Cherns [1]. He has kindly given us, as contributors to the books, an opportunity to comment on the review.

After listing the changes for the worker that follow from the introduction of semi-autonomous group working, Professor Eilon writes that it is "No wonder that 'group technology' has increased worker satisfaction and moral (sic), has reduced absenteeism, and has consequently lead (sic) to improvements in key productivity performance indicators." [2].

Granting this central point, he raises what he regards as two very important issues: how are we to measure the net benefit to the community? And how are we to assess the net damage to those individuals who do not find it easy to advance their personal ends in small working groups?

We regard these as important practical problems, not as theoretical dilemmas. We do feel, however, that practical problems have to be kept in perspective, particularly with respect to time. These volumes report the gathering in 1972 of a large number of people who had become concerned about the quality of work. Few indeed had been actively engaged in field experiments and the associated problems of measurement. Of those who had been actively engaged, none had had a chance to measure the effects in other areas of community life. Many of those who came from United States universities realized that their chances of being funded in this field of study were slim unless they could show that the results were measurable. We think the papers on measurement should be viewed in this perspective. They were of necessity exploratory; and probably, for just that reason, richer in ideas for the future.

Let us add but one observation. Measurement is not yet a major problem. In practice it is not difficult to find measures for changes in a particular plant or office that management, workers and unions can agree on. Getting measures that could be put in the annual company report to demonstrate social benefits is further away, but perhaps not so far away as Professor Eilon imagines; many people have remained quite active in this field since 1972.

Professor Eilon's second problem is again a practical one: How is the individual to be protected with the introduction of semi-autonomous group working? Most critically it should be noted that this form of working frees the individual worker from the petty tyranny of supervisors. They cannot be needled and treated as children as in the old way. We cannot bring ourselves to consider the supervisors as being more important than the workers. As management have depended so heavily on these people for so long they should be prepared when eliminating their role to give them fair reward for services rendered, What we do see as a problem is that some older workers may have become so accustomed to their narrow tasks that they want nothing else; However, they have lived with such tasks for decades and there seems little reason why they cannot live with them for a few extra years. Maybe there are people who for reasons of birth, family background, schooling and experience want nothing more of their work. While we suspect that this problem will be around for quite a while, we consider that it refers to a tiny minority.

We readily agree with Professor Eilon that the members of the conference of 1972 were enthusiastic but short of experience. We do not readily accept that workers or managers are so dependent on what social scientists accept as measures. No great precision is necessary, for example, before many significant decisions can be taken. Further we do not readily accept that workers are as hard-hearted about their peers as have been supervisors and their managers. We will allow one exception to this: when management set up a scheme of rewards that can only be achieved by the fit and greedy driving their peers to greater efforts. We certainly cannot agree with Professor Eilon that being motivated to improve things 'almost invariably' introduces the sort of bias in measurement that nullifies the scientific value of the research. Is it not possible, for instance, to do scientific research into cancer with the explicit bias of eradicating it? We can imagine that someone looking for a thesis topic might wander into this field with an attitude of sheer indifference. But our minds

boggle at the sort of misanthropy involved in starting from 'different premises' [2, p. 368].

REFERENCES

1. DAVIS LE and CHERNS AB (Eds.) (1975) *The Quality of Working Life*, Vols. I & II, The Free Press, New York.
2. EILON S (1976) (Editorial) The quality of working life. *Omega* **4**, 367–373.

FE EMERY

Centre for Continuing Education
The Australian National University
Canberra, A.C.T.

AW CLARK

Department of Sociology
La Trobe University
Bundoora, Melbourne
Victoria

Vol. 5
1977

OMEGA, The Int. Jl. of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

Strategies for Europe Proposals for Science and Technology Policies Industrial Innovation in Europe[1]

UMBERTO COLOMBO

General Manager, Research and Development Division, Montedison, Milan, Italy

(*Received May 1977*)

Europe's performance with respect to industrial innovation is compared to those of USA and Japan. The reasons for Europe's relatively poor performance are examined and related to institutional, cultural and social factors. An appropriate strategy for innovation in European industry is then defined.

THE PROBLEM OF INNOVATION

INDUSTRIAL innovation is one of a country's principal elements of vitality, and also one of the most important conditions for the success of its economy.

At the present stage of industrial development, innovation is of prime importance for Europe, since without it Europe's position would be weakened as a consequence of the structural changes that are taking place in a rapidly evolving world economy.

Innovation involves a whole series of factors, not only of a technical nature, but also productive and commercial, which must be viewed in the light of their relationship to economic, political and social conditions. In certain cases the adoption of technical innovation inevitably imposes parallel or prior social innovation. Moreover, an excessive social stability may hamper interesting technological innovations.

The elements that not only open the door to innovation but indeed stimulate it, are technical achievement through R&D, entrepreneurial thrust and, again, the attitude of society itself towards innovation.

A society such as exists in Europe, despite the wealth and variety of its

[1] This paper was presented at a Symposium organized by the Science Policy Foundation in collaboration with The Commission of the European Communities in Brussels from the 14–16 April 1977. Proceedings of this Symposium are to be published by Pergamon Press Ltd.

culture, and despite the turmoil, largely initiated by the new generation, has remained substantially anchored to its past and manifests fear of the new, by which it has been so frequently seared in recent years.

In these conditions, objective economic difficulty, due to the scarcity of financial means and national resources, has led Europe to an attitude of defence and self-protection, rather than to more aggressive and risk-taking behaviour.

A COMPARISON WITH THE USA AND JAPAN

For a deeper understanding of Europe's attitude towards innovation an analysis must be made of its causes. In this connection it might be interesting to draw a comparison with the behaviour of other industrialized societies, like those of the United States and Japan.

Europe has a great tradition of scientific research, at both experimental and theoretic levels, and has thus been and continues to be quite inventive. Europe lacks, however, that spirit of liveliness and entrepreneurial thrust, downstream of invention, necessary to diffuse the innovations on the market. By contrast, in the United States this spirit continues to be a true characteristic of the country, even if in recent times, there too, there are signs of diminishing innovation.

Vol. 5
1977

The United States is a young, enterprising and lively country with a marked propensity towards geographical and social mobility and with less rooted traditions than Europe, a country which looks to the future with hope and where the idea of growing material wealth for all has been until a short time ago, the foundation of life. In that country today, where the traditional views regarding wealth, economic growth and strictly material demand are subject to an evolution towards more qualitative objectives, there remains however a fundamental optimism, a confidence in the future that is a spur to innovation.

The legislative and educational systems in that country are structured in such a way as to favour and enhance the value of entrepreneurial drive. The laws favour a market economy and competition. We have only to consider such features as: the anti-trust law and the protection deriving therefrom for the smaller firms, which are thus in a position to compete with the more powerful companies; the disinvestment directives to the large corporations in order to favour free competition and prevent cartels and excessive concentrations, which on the contrary would appear to be natural in Europe; the antimonopoly action to which companies like the Standard Oil, du Pont, Western Electric, IBM, etc. have been and still are subjected, so that they would not become absolutely monopolistic in the oil, chemistry, telephone, and computer fields.

Its 'super-power' position imposes on the United States a major commitment to advanced research in strategic sectors such as, for example, aero-space, super computers, automation, non conventional materials. Progress in these

sectors induces further innovation in a number of other sectors of economy. But the most positive element that has up to now favoured innovation in the United States is the diffused belief in the superiority of the market economy, and the adaptation of the various factors (including the educational and the banking systems) to this belief. As a result, innovation in the US is of a strongly 'market-pulled' type.

If we look now at the situation and problems of Japan, it may be seen that the constraints of that country are similar to those of Europe. Japan, in fact, has serious shortages of raw materials and other resources (including space) and is involved with massive problems of an ecological and social nature (tensions, changing societal values, etc.). Japan, however, is not framed within a Community like the European countries, where each one seeks to optimize its own particular situation in relation to that of its neighbours. On this point, incidentally, it should be mentioned that the very efficient functioning of the Federal Republic of Germany, and its hegemonic economic position in Europe, is at least in part the result of the integration, due to the Common Market, of its economy with the weaker ones of its Community partners.

The geographically isolated position of Japan has caused that country, unlike Europe, to pay a great deal of attention to its relationship with those countries, including many of the less developed ones, with which it must keep up strong ties in order to assure steady supplies of raw materials, and other primary goods and to which Japan must in turn supply its products and technologies.

Japan, too, operates in a market economy, but in a rather different one from the American. It uses the domestic market, in fact, as a solid base for its productive expansion. Thus Japan practices a clear strategy towards foreign markets on which its industrial trading system behaves in a coordinated manner, unlike the United States for which market behaviour abroad follows substantially the same rules as are applied at home.

In Japan there is a close relationship between the political and the economic systems, with an interweaving of relations between Government, banks and other financial institutions, industrial and trade structures.

The economic and trade policies followed by the Japanese Government are chosen in close contact and coordination with industry and this, contrarily to the American and European approach, is not viewed as a fault, but as a positive feature. Similarly, the legislation of the country is devised to protect and to favour domestic industry.

After the last world war, Japan made precise strategic choices. Its military defeat induced Japan to emulate the victor, in a spirit of humbleness, once its ambition of domination was overcome. In the early post-war years, Japan thus copied as faithfully as possible external technologies and productive structures, particularly the American.

On this tissue Japan then grafted an innovation that has taken on an increasing

weight and depth, resulting from independent research, until achieving positions of leadership in certain, even highly important, sectors such as the iron and steel, shipbuilding, electronics, instrumentation, watch-making, photographic, motor cycle industries, etc. Thus Japan has gradually lost its inferiority complex, overcome its imitative strategy and, through its own inventiveness has been capable, indeed, of turning situations of disadvantage into big market opportunities, thanks largely to industrial innovation.

The innovative capacity of Japan has been particularly apparent at the organizational and job efficiency level, in the light of a systemic approach that allows coordinated action, a kind of team game, of the various forces (political, financial, productive, commercial and social), each of which respects the role of the others. This approach therefore acts as a multiplier bonded to the structure of society.

This unusual innovative performance is to be related to the specific culture of Japan and, possibly, also to the use of an ideological language and of a corresponding ideographic script. The language for ideas, in fact, favours a software, rather than a hardware, approach, and a probing in the direction of the function for which something is thought of or produced, rather than in the direction of the specific technology currently available for its attainment as often occurs in Europe, or of the market that is to use it, as frequently is the case in the United States.

Vol. 5
1977

The Japanese tendency towards a highly precise vision of things and towards miniaturized work, has initially favoured the penetration of Japan in such technologies as precision mechanics and such industrial sectors as watchmaking. The Japanese way of thinking, however, less bound to the type of technology and more directed towards the function of use, has led the Japanese industry to move forward to other technologies, more suited to fulfill the function of time measurement. The Japanese watchmaking industry is thus gradually disengaging from the traditional technology and adopting non conventional solutions, such as quartz oscillators, optoelectronics, etc. without the psychological attitude of a follower with respect to the leading Swiss industry.

The Japanese picture, with reference to innovation, is also correlated to the very high average level of education. The strong innovative energy characterizing the Japan of today should also be referred to the scale of values of the Japanese society, partly still highly traditional and partly incredibly modern. These values are allied to a marked '*esprit de corps*', that goes from the enterprise level, up to a deep-felt spirit of national solidarity. For Japan this is a necessity of life. To this must be added a strong desire for revenge, which however from the military sphere has been transposed to the economic sphere (production and trade).

The situation of the United States and that of Japan, in perspective, are certainly not immune from danger. In the United States, in fact, where Government intervention is by many viewed with diffidence, the rôle of Government in

organizing social demand has been relatively minor, while industrial activity has prospered on the remarkable mobility of the labour force and on a comparatively high level of unemployment. All these elements have been decisive for the development of the American economy, but the values lying at their base are unlikely to be sustainable. In Japan, instead, the still low but growing social and generational tensions may well loosen the tight knots of solidarity and accentuate the workers' claims. Neither must it be ignored that Japan has operated as the center of a vast geoeconomic system, to which it provides goods and services in exchange for resources and raw materials, whose price has increased considerably in latter years. Furthermore, Japanese exports include many manufactured articles that call for a considerable availability of labour, whose cost is no longer so low in Japan. This situation, too, is destined to further deterioration in the future.

If we now examine the context and the approach of Europe in relation to innovation, it has to be acknowledged that other factors have acted in a negative sense, in addition to those previously mentioned. The great and diversified cultural wealth of Europe has not so far succeeded in becoming a valid instrument for innovation. Europe, moreover, has been conditioned by its history of glory and power, by its political and cultural heritage, by its strong traditions, which are hard to leave behind, whilst at the European level there has been no lesson learnt from a lost war, apart from the case of Germany, that represents an occasion for national or cultural reassessment.

The positive approach towards the market economy and towards the new, typical of America, is far less followed in Europe, with its experience of colonialism and of privileged markets. This experience has brought forward a pronounced hierarchy of economic, social and cultural values.

European tradition thus leads to a greater social rigidity as compared with the United States. The 'social climbing' phenomenon is still much less common in Europe, whilst family tradition, name, the fact of belonging to certain groups continues to have considerable importance.

In Europe, laws and public attitude have not favoured lively and diffused competition, but rather have permitted the formation of monopolistic or oligopolistic situations in several industrial sectors. Similarly, industrial cartels of a defensive character and therefore substantially non-innovative have been formed contrary to what takes place in the Japanese industrial and trading concentrations, whose function is, instead, directed towards attack abroad.

In European conditions, the market proves to be considerably less stimulating for the diffusion of innovation than in the case of the United States or Japan, whilst the strong scientific and technological tradition causes the 'technology-pushed' type of innovation to be more important. This appears obvious when consideration is given to the sectors in which Europe is strong, such as precision mechanics, optics, machine tools, pharmaceuticals, dyestuffs and other fine chemicals.

Examples of technology-pushed innovation occurring more recently in Europe are the float-glass process and the L-D oxygen steelmaking process, which have spread throughout the world; air transport by helicopters, which however later found a greater independent development in the United States; sea transport by hovercraft, which has not yet attained a full market success; and, at the product level, polypropylene, from which the company that invented it has not obtained the best advantages in the marketplace; or carbon fibres, that represent an unquestionable technological achievement, even if the innovation is diffusing in the market at a very slow rate.

This technological ability of Europe allows, in some sectors, the positions of strength attained through innovation to be maintained (as in the case of pharmaceuticals, dyestuffs and other fine chemicals), but Europe is not always capable of fully exploiting the potential of market development allowed by technological innovation and sometimes, is not even in a condition to defend positions of initial strength, against continued innovative competition from abroad.

This situation is aggravated by the fact that the lack of political unity does not allow the concentration of effort and the strengthening of technological competence necessary to cope with dangerous situations in an innovative manner. This is the case, for example, of the Swiss watch-making industry, which had certainly been able to foresee threatening developments due at first to large-scale production and to the consequent mass approach on the market, and later to the introduction of electronics. However the Swiss watch-making industry was not promptly in a position itself to put these developments to use because of an insufficient blend of internal skills. The Swiss watch-makers are now catching up with this disadvantage, and to this purpose are lining up with the new technologies, and integrating their previously missing skills. The extent of the success will depend on the capacity for asserting, together with the adoption of the new technologies, the value of the Swiss name, which traditionally is still a guarantee of competence and capability. Perhaps, however, the major effort of technological and managerial adaptation that the Swiss watch-making industry is undertaking would lead to more rapid and effective results if the Swiss industry were better integrated within a European framework.

Europe often ends up by tailing after innovation, the thrust of which is given by others even when, in some cases, the invention originated in Europe, and doing this it is more difficult to be competitive. This is true for some branches of electronics and informatics, for precision mechanics, for scientific and control equipment, sensors, etc.

It is important, therefore, to better understand why Europe is often more inclined to operate in sectors where the continued thrust of innovation comes from other quarters, adopting a defensive attitude in relation to the aggressive posture of the countries with which it has to compete.

For this purpose it may be useful to refer to other sectors, in addition to those

Vol. 5
1977

previously mentioned, such as shipbuilding (a sector of undeniable European origin), the heavy steel industry and the automotive industry. In these sectors Europe has had to yield ground to Japan to an increasing extent, not so much because of strictly technological deficiencies (even if Japan has realized interesting technological innovations, like the new ship-launching techniques, or the lowering of energy consumption in the large scale iron and steel industry), but rather because of the great ability of the Japanese in work organization, and in their general vision of production and commercial creativities.

To this must be added the capability of adaptation of the Japanese worker who, in common with his American counterpart, is far more willing than the European worker to accept burdensome and hard working conditions. However, the Japanese society, of the patriarchal type, supports the worker to the utmost, whereas the American—decidedly more individualistic—offers the worker opportunities to find a new job and incentives to stimulate his commitment. Europe, instead, has a long tradition of social protection of the worker—a tradition that represents an essential value for its society—and the worker himself tends to become increasingly involved in the decision-making process through phenomena such as industrial democracy and 'mitbestimmung'. This different situation may lead, at least temporarily, to lower competitiveness of European industry.

All this still does not represent a complete explanation of the unsatisfactory innovative capacity of Europe and its tendency merely to follow up the innovative process led by others.

The technological approach to innovation, typical of Europe, has caused a neglect of other types of innovation, for example, innovation in marketing. In this connection, mention may be made of the Xerox process, which while being connected to an unquestionably major technological invention (the importance of which was promptly understood by the British), was favoured by an equally big marketing innovation: namely leasing the copying machine. This latter innovation has brought into being a new mode of use, switching the cost to the consumers from initial investment to day-by-day use. This type of innovation, which is based on the analysis of latent demand and which, in the case of Xerox, has caused mass communications to explode in the business area, is typical of the USA, which is remarkably prompt at anticipating market needs, integrating technological innovation with marketing innovation, thus allowing the former to develop.

This readiness to innovate in relation to market needs has, however, aspects that should be considered critically in that it could favour a propensity towards a 'throw away' society and planned obsolescence, that is to attitudes that might instigate crises in that such attitudes are becoming increasingly less welcomed to a more attentive and informed society.

Furthermore, the European market cannot realistically be regarded, for European industries, a true 'domestic' sales outlet. The Common Market is

certainly not such, as far as public sector purchases are concerned, and in Europe the 'buy European' issue is practically non-existent. The American and Japanese markets are, instead, well protected and isolated; they actually represent big domestic markets readily accessible to the respective industries and, moreover, as far as public purchases are concerned, there is more co-ordination and certainly not the type of bitter conflict that exists among the European governments that tend to lean towards their respective national industries.

A STRATEGY FOR INNOVATION IN EUROPE

To enable the European situation to be modified in a positive sense, a clear vision would be needed of the objectives intended to be pursued and of the means necessary to achieve them. This implies the adoption of a strategy that Europe has not thus far been capable of providing.

The problem is quite different for the United States, which is one of the two world leading powers, with the structure and strategy typical of leaders. The strategy of the others must be modelled on that of the leader and, according to economic circumstances can be a strategy of defence, or of attack, but always a strategy of coexistence.

Vol. 5
1977

In the immediate post-war years Europe was engaged in reconstruction with the benefit of American aid which as a consequence favoured the adoption of the American model in Europe. This kind of dependence should be viewed in relation to the pattern of technology adopted which was quite different from the one that Europe had been independently developing up to that time. With the acceptance of a model of production based on large scale units and parcellized work, a certain degree of downgrading of the European labour force has also occurred.

To this Europe has added over time a series of endeavours, such as the supercomputer, the supersonic civil aircraft, and the fast breeder nuclear reactor, thus pursuing agelong dreams of glory and power. Big projects were consequently conceived and organized. The lack of political unity and the divergence of economic interest among European countries, have been major factors of failure in the first two instances. The fact that Europe is not a political superpower may be detrimental to the third of the above mentioned projects.

These examples indicate that in such cases it is not sufficient for Europe merely to have its technological papers in order.

To achieve a stable position of competitiveness, Europe must give itself a strategy, and be ready to adjust it to the rapid evolution of the world economic and socio-political environment.

Europe, as we have seen, is in a considerably more complex situation than is Japan. In the first place, Europe is not politically united, and the economic

integration process started in the 50's has so far not been pursued to a sufficient extent. In other words, Europe is still facing the problem of how each of its member countries may fit itself into a framework of progressive European economic integration, and this eludes the more important problem of how to optimize the situation of Europe as a whole, in a world economic context.

This means that Europe has been incapable of giving itself a unitary strategy, and that the process of strategic planning (if I may be allowed to draw a comparison with the industrial world) is more similar to that of a loose conglomerate than to that of an integrated and well coordinated industrial enterprise.

We certainly cannot wait for the event of a European political union to arrive at a European economic strategy, and neither, for that matter, would the setting up of a unity at a political level be in itself a sufficient premise for a valid strategy. What, however, needs to be done is that in tracing the outline for even a rough and limited converging strategy of the different European countries, account be taken of the objective of a longer term political unity. In other words, it must be assured that a strategy of economic and social development in the different European countries does not hamper the indispensable objective of integration also at the political level.

Since Europe, as has already been mentioned, has a dramatic scarcity of natural resources, it is clear that its economic strategy and its very social welfare depend to a notable extent on the competitiveness of the industrial system, and this in turn is conditioned by innovative capabilities. It is obviously not feasible for each individual European country to innovate on its own account along a broad front. In fact, there already exists, substantially, a more or less marked specialization of the different countries in different technologies and industrial sectors. As things stand at present, these specializations cannot be rationalized so as to produce an accentuated regional division of productive activities at a European level, as would be easier to achieve in a Europe with united political and monetary interests.

Looking at the situation of Europe in relation to its interdependence with the other countries of the world, it is unthinkable that the process of an international division of labour could be allowed to proceed to the point of accepting a total dependency on outside supplies, for instance, of base-strategic production such as iron and steel, non-ferrous metals and other important materials, fertilizers, fibres, etc.

One of the fundamental elements, at least for the coming decades, of a European strategy must therefore be the rationalization of the base industries, which are characterized by a relatively low unit value of the products, and by processes that tend to be increasingly capital-intensive and concentrated in a few large plants. Rationalization in these base industries should lead to a better balance between centralized and decentralized production, the latter capable of compensating the apparent diseconomy of scale with a greater flexibility in production and hence a greater adaptability to market demands. A further

element of rationalization is the distribution amongst the different European countries of these base industries generally classified among 'mature' sectors. A third and perhaps more important element is the degree of dependence on imports which, for such sectors, Europe is prepared to and can accept. This problem is particularly relevant to the issue of industrialization of the less developed countries.

Besides the base industries, there exist a number of industrial sectors for which one or another of the European countries is in a leading position at the world level. Conversely, each country is weak in other sectors, with respect to which it has often in the past followed policies of a protectionist type, safeguarding, that is, its own position even with no pretense of converting such sectors of weakness into sectors of strength. Viewed from a European standpoint, this policy of assistance to weak industries certainly constitutes a strategic error, since it tends to disperse intellectual, entrepreneurial and financial resources over a too broad arc of activities. A greater market specialization, at a national level, can and must be achieved in Europe, and perhaps it is easier to start from the less critical sectors. This specialization should be favoured by technological transfusions from one sector to the other, and from one country to the other, so as to increase the competitiveness of the different European countries in the respectively chosen sectors, at a world level. Such a strategy should also be applied to those advanced sectors which do not call for an exceptional commitment of human and financial resources. When, instead, such a commitment is necessary, Europe will have to move from the outset in an integrated way, with a highly careful choice of the sectors in which to commit itself, avoiding the ones that are motivated essentially by considerations of prestige and not backed by solid economic prospects, that only a united Europe, with its major political weight, could sustain.

Vol. 5
1977

Going back to the more general problem of Europe's rôle in technologically advanced sectors, the problem is not so much that of Europe's presence or absence in a given sector but, rather, that of the type of its presence and of the related attitude. Europe, for instance, is certainly capable of developing its own strategy in the sectors of informatics, telecommunications and space technologies without necessarily following the models set up by others, and therefore without yielding to the temptation of committing itself to a 'follow-the-leader' type of research which most likely would produce negative results.

In the overall context of industrial policy it is in any event clear that even an economically integrated Europe would not be in a position to maintain a valid competitiveness in too many sectors. When strategic considerations related to the safe and continued availability of indispensable goods do not prevail, Europe should decide upon and willingly accept a substantial dependence on other countries. This type of strategy becomes a necessity if we wish to resuscitate in the European context the positive elements of a market economy, namely efficiency, entrepreneurial spirit and innovative thrust.

It must be acknowledged that in Europe a mistaken policy of protection of weak sectors or, in other cases, the predominance of oligopolistic forces bringing about the formation of producer cartels with assured profits for all members, have led to a dulling of entrepreneurial spirit and to a de factu limiting of innovation. This situation must be firmly rectified by enlivening competition within Europe and possibly by joining the interests of European producers regarding export policies, following, in this, the example of Japan.

In cases where it might be necessary to temporarily retain cartels or other forms of concentration among producers in a sector (for example an industry of strategic importance in a state of crisis as to-day is the case of man-made fibers) a rôle of control and supervision by Governments and preferably by the Community itself would appear indispensable.

Europe had a long tradition of sophisticated production that called for highly complex individual skills. This tradition, at least in some of the European countries, has flagged in the latter decades for a series of reasons, in parallel with the occurrence of a process of concentration of production in larger plants. This trend should be reversed, and this could be achieved through a new decentralizing thrust, connected on the one hand to a better regional management including the safeguarding of the European cultural heritage and, on the other, to the necessity of reducing unemployment.

The requirement, previously referred to, of reinforcing in Europe the positive elements typical of a market economy must not give rise to a push towards a 'laissez-faire' type of economy that would lead to increasing disparities and internal contradictions.

On the other hand, a rigidly planned economy discourages an innovative attitude. This can only be nurtured through greater participation and involvement at all levels. Rigid planning results in a bureaucratic 'top-down' rather than in a 'bottom-up' approach, thus hampering entrepreneurial spirit.

It is therefore necessary to attain a harmonious balance between the characteristics of a free-market and others of a guided economy.

THE EUROPEAN ROLE IN SOCIALLY ORIENTED SECTORS

The European tradition of labour protection, its policies of low unemployment, of the welfare state and related social infrastructures, its performance in urban planning (for example, in the case of the British 'new towns'), its vanguard position in promoting a different and less alienating way of producing in industry (Volvo for instance, but also Philips, Olivetti, etc.), the close relationship between Government and industry, and the extent of direct political intervention in the national economy to achieve a better adaptation of industry to

social requirements, all these elements should place Europe in a favourable position vis à vis socially oriented innovation, the need for which is strongly felt in modern societies.

The increasing interaction between agriculture, industry and services, which is common to all advanced societies, is taking place to a considerable extent in Europe, which is pressed by such problems as the scarcity of natural resources and high population density. This process of gradual convergence of the productive and service structures offers industry the opportunity of playing a primary rôle in the fulfilment of important social objectives. The innovative capacity of industry, its organization—flexible enough to face different and complex situations—make it a suitable instrument to cope with the problems posed by a highly interconnected system of activities. The facts that the direct goal is the satisfaction of social demands and that the related expenditures are, substantially, part of the public budget, make it necessary for governments and public administration to control and guide the action of industry in these sectors. Industry must understand this function of the public power, and cooperate without pretending to be capable of doing everything by itself.

It should be realized that these socially-oriented sectors expressed by public demand cover a very wide range of activities and involve problems that are by no means trivial. Under this aspect it is once again necessary to reflect on the potential benefits that Europe could obtain by adopting a coordinated approach, that is by gradually replacing the public demands of the individual countries with a much stronger European public demand. This would allow a broadening of the market base for the related goods and services, and would favour a higher degree of specialization and professional content—hence of innovation—of the industries concerned.

Vol. 5
1977

It is clear that to give the state a more important rôle as a purchaser of industrial goods and services means to reduce—*ceteris paribus*—the volume of expenditures that the citizen can afford on a private basis. This should entail a more austere way of life. The reduction of the level of private expenditures could, however, prove to be less dramatic and smaller than might be thought: in fact, the deprivation of potentially available private goods is at least in part balanced by the reduction of social costs deriving from the new mix of economic activities. Moreover, this deprivation may be more easily accepted if it concerns products, which have given rise to a high level of waste of resources.

Furthermore, one should not neglect that these socially-oriented technologies and products are indeed exportable either as such or after adaptation to different situations and needs. There is, in fact, a growing demand—particularly in the developing countries—for goods and services related to social infrastructures (schools, hospitals, public transport, housing, etc.), and this could provide Europe with an additional opportunity of income and of contributing to socio-economic world development. It is, however, important for Europe to maintain a dynamic pace, if the potential advantage of its higher propensity towards

socially oriented innovation is not to be lost in favour of more aggressive and market-oriented competitors.

An analysis of socially oriented industrial activities leads to the consideration of a direction followed so far by industrial innovation in general: that is, towards the expulsion of labour from the productive process. In fact, the concept of an ever increasing productivity of labour, favoured by technological progress and innovation, has been traditionally associated with the idea of an expanding and growing economy. In conditions of scarcity of financial means and natural resources, the assumption of an ever growing economy, where the resource represented by labour is relatively scarce, is seriously challenged. If even during periods of sustained economic growth the level of unemployment has increased in industrialized countries as a result of continuing productivity increments, this situation could, at least in the next decade, become more serious in view of a slower rate of growth of the economy. A crucial issue in this respect is the understanding of the relationship between innovation, from the initial discovery to its wide diffusion in the market, and unemployment.

As indicated by Freeman, in the initial stages of innovation—associated with key patents, many prototypes, intensive R&D for new products and applications, little standardization and relatively small scale production—jobs are created as a result of the innovative process. In contrast, in the later stages of innovation when the goals of R&D have shifted largely to cost-saving, most applications have already been developed, products are largely standardized and manufactured on a large scale, labour is displaced.

While it is historically true that over the lifetime of a basic innovation more jobs may be created than displaced, the problem has different connotations when considered in the light of the international economic context. It is clear that those countries that do not contribute sufficiently in the earlier stages of basic innovations are forced to import unemployment, in the hidden form of the final stages of innovation. Under this aspect, innovation, called upon to solve certain socio-economic problems, may in the end create new ones in the form of unemployment. This is clearly an oversimplified discussion of the impact of innovation on employment: for example, the relationship between capital and labour, the degree of their substitutability or complementarity should be further investigated and included in the picture.

Technology should in the future become 'appropriate' to the changes in factor proportions (at least in the medium term, surplus of labour and scarcity of capital). This change of factor proportions could serve in itself as a stimulus for new and different innovations, as opposed to artificially, if not forcefully, increasing labour intensity in existing activities where a surplus of labour leads to uneconomic results.

This leads to the problem of identifying and developing technologies 'appropriate' to the European economic and social system, where the abundance of

labour is not compensated by a low unit cost of labour. Several industrial sectors lend themselves to the development of 'appropriate' technologies economically viable in a European context. For example, the utilization of local mineral resources to manufacture building materials; the reclaiming of scrap, the utilization of waste for different purposes including the production of energy; the decentralized energy systems utilizing various sources including the minor ones, and exploiting in an integral and optimized way the different forms of energy produced (total energy systems); the introduction of biological concepts and methods within chemical and agricultural production; and so forth. The development of appropriate technologies in all these sectors could lead to productive activities characterized by higher labour and lower capital intensities.

Needless to say, these appropriate technologies must not be seen as counterposed to the traditional ones, but rather as complementary to them in the context of a better balance between centralized, capital intensive and decentralized, labour-intensive activities. The real objective is to pursue an appropriate combination of technologies within a strategy based on what may be called 'technological pluralism'.

Europe's great shortage of resources may be conveniently faced by considering, in addition to the use of conventional raw materials, often imported, the employment of alternative local resources, of substitute and more plentiful materials, of a developed recycle economy, of a different way of designing and of producing goods so as to lengthen their life and allow ready replacements and repairs.

Appropriate technologies and technological pluralism thus represent a complex and articulate pattern of production which must be conceived and understood at a system level. In this sense, they provide a further notable opportunity for innovation and entrepreneurship. They also appear as an opportunity for exporting technologies more suitable, after due adaptation, to the real needs of the developing countries.

Vol. 5
1977

SOME TENTATIVE RECOMMENDATIONS

We have so far outlined the elements of a possible European strategy in which industrial innovation would play a major rôle. The creation of a climate favourable to industrial innovation is a real issue, an issue that was analyzed several years ago within OECD. The Pavitt report (1971) that presented the conclusions of such a study suggested the following objectives for a governmental policy in favour of innovation:

- ensuring industrial competition, as the main pressure for technological innovation;

- ensuring equitable rewards for innovations, through the tax and patent systems;
- ensuring that regulations, codes and standards take account of both the social costs and benefits of the innovative process, as well as the flexibility and pluralism required for successful innovation;
- having active regional and manpower policies to deal with the changes in industrial and skill patterns brought about by technological change;
- using government procurement to upgrade the technical level of industry, and to couple technology more effectively to collective, social needs;
- encouraging the mobility of scientists and engineers, especially in and out of government laboratories;
- identifying policy measures to encourage science-based entrepreneurship;
- ensuring continued trade and capital liberalisation, thereby heightening the pressures and incentives for technological innovation in all Member countries, and maintaining the rapid, international spread of the benefits of new technology.

Most of these objectives would be valuable now to implement a European strategy for innovation. However, it is fair to say that the present international economic situation is more complex and difficult than when the OECD report was written. In particular, the financial difficulties in which many industrialized countries find themselves have led to a weakening of the market economy and to a more direct and determinant rôle for Governments. Therefore more emphasis should now be placed on the creation of active policy instruments, capable not only of promoting actions in the direction of public and socially oriented activities, but, even more important, of adding vigour to the spontaneous entrepreneurial forces of the economic system. Such instruments should on the one hand stimulate and adequately reward innovation, and on the other discourage and punish those enterprises which base their success on the creation of barriers limiting active competition and unduly exploiting the market.

In the case of industrial cartels, their elimination must take place in the context of a clear industrial policy at a European level. If, for strategic reasons, a sector has to be protected, this is not a good reason for avoiding intra-European competition. Rather, the protective measures must be adopted at the boundaries of Europe, and be motivated by serious and defensible arguments. This would put an end to the present system of aid to weak industrial sectors that each country carries on with its own criteria, and in some cases with a short-sighted approach.

A 'Buy-European' policy is needed on the part of the Governments of the various countries, particularly in the technologically advanced and in socially oriented sectors. The Government as a purchaser must not adopt a bureaucratic approach, but should instead aim at product quality and efficiency, thus stimulating innovation.

The medium and small size industries have traditionally played an essential

rôle in the generation and diffusion of innovation. Adequate instruments to support the action of the innovative smaller enterprises must be devised. Today instruments aiming at this objective are operating at a national level in several European countries (ANVAR, INOVA, RKW, and also, to a certain extent, NR & DC, the IMI Fund, etc.). So far there is no instrument of this kind available at a genuine European level. A venture capital enterprise (EED) formed about 10 years ago on a European basis following the example of a successful practice in the United States, has failed in its efforts largely because of the traditional attitude towards risk of the financial stockholders. A more recent initiative in this direction, SCIENTA, is perhaps in a better shape, although one cannot refer to it as an ongoing success. In any case, there seems to be a need for a European venture capital enterprise, managed with an entrepreneurial, risk-taking attitude and with adequate technological foresight. If the European banking system is not available for such an initiative, and if the large industrial concerns are not interested in being the stockholders of such an enterprise, the EEC could exert promotional action and perhaps even sponsor the undertaking at its early stages.

The big industrial companies should also catalyze and sustain innovation within the smaller firms, by putting into being laboratories or experimental centres for the transfer of technology, aiming at providing technical assistance and at stimulating the innovative effort of the assisted firms. Such laboratories could receive a public incentive, in view of their non strictly promotional scope. Similarly, government encouragement should be given to the professionally worthy sectoral research and development associations, which are often the cooperative effort of many firms of smaller size.

Vol. 5
1977

As far as large innovative projects are concerned, often requiring major scientific and technological effort and a heavy financial commitment, a selective and attentive policy should be set up at the Community level. This case is but one instance of the usefulness of installing a European Office of Technology Assessment, as suggested by Europe plus 30.

Among actions that would certainly contribute towards stimulating innovation in Europe, those favouring trans-sectoral and international mobility of skilled individuals should not be overlooked. This issue of mobility is one of the main elements of a sound manpower planning policy.

One last comment: industrial innovation is the product of the intellectual effort of capable individuals. In latter years the attitude of people, and of the young in particular, towards science, technology and industry has become more critical and suspicious. In fact, much of the unrest and difficulties of our time is ascribed, without a critical analysis, to science and technology, to the point of assuming in certain cases an '*a priori*' attitude of total rejection of technological innovation.

Industry in the past has taken many decisions without a sufficient sense of social responsibility, considering that its main duty was to provide an increasing

wealth to society. In doing this, industry has sometimes exploited the lack of knowledge, within public opinion and administration itself, of possible unwanted side-effects of industrial production or products.

Now industry must realize that a process of technology assessment should be incorporated as early as possible in its strategic decisions, and that a frank dialogue with the educational system and public opinion is in its own interest. This dialogue will provide a better understanding of industry's viewpoint and of the role that science and technology are now, more than ever, called on to play in answer to today's problems. Action of this kind should also favour the recruiting, for industrial research, of young individuals, endowed with the necessary ability and motivation.

These and other general recommendations, in order to be implemented, need a thorough evaluation of their institutional feasibility and of the likely consequences of their implementation. It seems therefore advisable, if the EEC Commission is to play a concrete rôle in this issue, that a project on industrial innovation in Europe be launched and that close links be established with the directorates dealing with scientific and technological policy, and with industrial and economic affairs.

ADDRESS FOR CORRESPONDENCE: *Professor U Colombo, General Manager, Research and Development Division, Montedison, Largo Donegani* 1/2, 20121 *Milan, Italy.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977, Pergamon Press. Printed in Great Britain

Accuracy in the Long Run—Where Are We Now?[1]

HSD COLE

The University of Sussex

(*Received February 1977*)

Long-term forecasting must be viewed as informative speculation about the future. It should be credited with relatively little scientific authenticity. At the present stage of theory and data, no magic methods can be expected to overcome the problem of satisfactory forecasting. Improvement, however, is certainly not just a question of putting more and more variables and more and more numbers into a computer. Methods which indicate how to cushion against uncertainty and methods which bring a greater awareness of options for the future are essential if we are to gain greater control over events. The hazy images of the long-term future which are generated by scenario and other forms of analysis form the guidelines within which short and especially medium-term choices must be made. But in the end we must recognise the inherent limitations of forecasting and forecasting methods and think of forecasting not so much as a method of prediction but as a contribution to tackling the future in a more integrated sense.

WHAT IS LONG-TERM FORECASTING?

IN THIS paper I propose to take a questioning look at forecasting, in particular long-term forecasting methods. Essentially, what I will examine here is the possibility of providing a systematic explanation of the factors determining long-term growth rates in order to provide the background assumptions entering into the planning procedure. Certainly there are many people calling for this and many other people striving to provide it. I will confine myself to the broader aspects of long-term forecasting since examples of particular methods are to be discussed in later papers. If I am sometimes a little sceptical of forecasting techniques, including many of those advanced elsewhere in this journal (and therefore perhaps somewhat heretical in a place where quantitative forecasting is often assumed to be useful), this is all in an attempt to be con-

[1] In '*Forecasting for Business Decisions: A Collection of Research Papers*' (Eds. FILDES R and WOOD D) Saxon House, 1977.

structive about forecasting and how we should go about it. Furthermore, the general argument I am putting forward, as many of my examples will demonstrate, applies to short-term as well as to long-term forecasting [8, 4].

However, given that I am mainly concerned with long-term forecasting, how should this be defined to distinguish it from other types of forecasting? There is no straightforward answer to this but the simplest seems to be that time horizons are set by the characteristic time scales of the physical, social, economic and technological processes with which one is concerned. Thus, in short-term forecasting most factors may be assumed to be fixed, for example the structure of political systems and other social and economic relationships, and this is implicit in most of the methods developed for short-term forecasting. Long-term forecasting on the other hand usually takes the view that the future is more open to the effects of human action. In addition to this, however, many more uncertainties arise from the impact of one aspect of the world on another. In fact, because the further ahead one looks into the future the more factors seem to be implicated in any one forecast, almost any question eventually opens into a global context.

Nevertheless, this kind of separation between short and long-term forecasting may be of restricted practical use. If you have little control over events affecting your life or your firm, even the most short-term problem soon becomes 'long-term' according to this definition. Consequently I make a somewhat arbitrary choice (which is relevant to our own work at the Science Policy Research Unit) and treat long-term as anything between say 20 and 50 years, the typical lead time for a major research and development project to have a great social impact or for major shifts in world development to have taken place. This time period may not at first sight be directly relevant to the economic success of day to day industrial decisions but it is highly relevant to their collective social impact. Therefore, despite my scepticism about many attempts at forecasting, I believe it is extremely important that we should create a long-term perspective, in the hope of influencing our long-term future.

Vol. 5
1977

Forecasting over this time span cannot be defined in a narrow sense as an attempt to 'predict' the future. One can think instead of forecasting as consisting of three related activities. The first of these is a visionary activity in which one attempts to create images of the future, images which may be utopian or distopian. Traditionally these have been in the form of literary works such as moral and political writings and have often set out to question the basic values underlying human society. The future which is postulated thereby may be in a systematic fashion or else certain aspects may be described in great detail.

The second activity in forecasting is that of 'projection' in which the logic of the desirable, conceivable future is demonstrated. This is usually done through historical analogy of one form or another. It may be done on the basis of extrapolation, using more sophisticated mathematical models, or through the selection of telling historical events. Essentially what we attempt to do here is to

provide some theoretical understanding to back up a long-term view of the future. I would argue, however, that extrapolative methods as such are almost atheoretical in that little or no underlying formal theory is assumed although continuity is implied.

There are always trade-offs to be made between one aspect of the future and another and this leads to the third activity in forecasting, that of policy analysis. In this the repercussions and trade-offs are examined in some detail; there are methods such as cross-impact analysis and cost benefit analysis which have been developed in an attempt to solve such problems. Whether one begins the forecasting activity with the vision, the projection or the policy analysis depends on the particular problem that is being approached. However, in general, all three elements will be present in any 'forecasting' activity. I shall argue in the latter part of this essay, that the emphasis placed on each of these stages is often wrong. Too many hopes are placed on the second—the projection.

WHAT CAN BE ACHIEVED?

Recognition of the great complexity of long-term forecasting and at the same time the concern that there is a need for lengthening the time scale of accurate forecasting has led to a search for better ways of understanding and dealing with the long-term future. Ideally, one would like to achieve methods of forecasting which would help to identify:

(i) opportunities and dangers (arising out of present trends);
(ii) guiding context within which alternatives for the immediate future can be considered;
(iii) ways of making sure that the right choices are made now and that the right choices will be able to be made in the future (i.e. one does not want to foreclose too many options).

We also need to ask the question of what can be reasonably achieved. A great variety of methods have been put forward but what do these methods have to offer? In the first place they offer a greater level of rigour and confidence (in other words they permit us to make best use of available knowledge). In the second place they permit us, at least in some cases, to experiment with and evaluate the effects of alternative policy strategies. In the third, they improve communications between researchers and policy makers and the public at large. However, they only offer each of these to a limited degree and, for example, do not even necessarily lead to an increase in the accuracy of prediction.

I do not mean to convey the impression here that long-term forecasts aimed at helping to clarify public or commercial choices need be infinitely accurate. For many purposes quite rudimentary forecasts suffice, provided they lie

between the thresholds which dictate that a qualitatively different allocation of resources should be made. Nevertheless, this argument leads to the conclusion that in attempting to estimate the benefits of different forecasting methods we should of course have a thorough understanding of just how valid data and theory are in each of the areas where explicit or implicit assumptions are being made.

Immediately we come to consider long-term forecasting in any sector we realise how inter-twined these things really are. Nevertheless, because we can identify sectors of understanding often represented by the traditional disciplines it is instructive to consider how well we can forecast in each discipline separately and where the deficiencies lie.

Population growth

In demographic forecasting, for example, although there is obviously a great deal of momentum in the growth of population, our ability to forecast breaks down fairly rapidly through lack of understanding about, for instance, fertility rates which are known to depend in some way on socio-economic factors, but precisely how we do not know. Even over quite a short period changes in the population would depend on a variety of physical, environmental, technological, economic, social and political factors. [3]

Vol. 5
1977

It is hardly surprising, therefore, that in the field of demographic forecasting there have been a number of unexpected failures. For example, the UK Commission on Population some years ago carried out an extensive and sophisticated forecast of UK population. Many alternative forecasts were given to allow for foreseeable contingencies. Twenty years later the actual size of the population lies outside any of the possibilities considered. Forecasting 40 years ahead there is at least ten millions uncertainty in the likely population of the UK (it should be remembered that this is in a country with one of the longest histories in demographic forecasting). Table 1 shows the variation between official

TABLE 1. VARIATIONS BETWEEN THE OFFICIAL PROJECTIONS MADE IN 1953, 1958, 1963, 1968 AND 1973

Year from which projection was made	Projected total populations of England and Wales (millions) 1963	1973	1983	1993	2003
1953	45	46	47	46	46
1958	46	48	50	51	53
1963	47*	51	55	59	65
1968	—	50	53	57	62
1973	—	49*	50	51	53

* Actual.
Source: Cox and Wilson [7].

projections, to the turn of the century, in years between 1953 and 1973. Errors in the forecasts of migration have been larger but the total numbers involved have been small in relation to the national population. In the main, however, forecasts of working-age and retirement-age populations have been sufficiently accurate for most plans except the most long-term [7].

Climate

One discipline rarely taken into account in long-term forecasting is the study of weather and climate despite its potential importance for agriculture. Even in the short term and despite the use of very sophisticated mathematical models, and methods of obtaining data, we are hardly beyond the stage of being able to forecast day to day weather beyond the old rule of thumb of saying that tomorrow's weather will be very much like today's. Fortunately, this situation is slowly improving. With regard to forecasts of long-term climatic changes (that is the prediction of local or global weather shifts or the forecasting of ice-ages), which may be of more importance to the long-term forecasting of energy requirements, for example, we are in a much worse situation. More or less equally reputable climatologists are forecasting the advent of a new ice age or (as a result of man's activity) a melting of the ice caps within the next century [12]. We consider other aspects of the physical sciences later.

Consumption of physical resources

An area directly concerned with long term forecasting is the question of long-term availability of physical resources. Here we run into problems of definition. What is a physical resource? In the long-term it must have a social and techno-economic definition [8]. Resources important a century ago, for example wood, are now of a much lesser importance and resources at that time hardly used such as oil are now dominant. Attempts in the past to forecast long-term resource availability (e.g. the Paley Commission on Resources) have nearly always fallen down because of the uncertainty about the actual amount of the material which would be made available [18].

Consideration of this report in somewhat more detail here will give some idea of how well longer-term economic forecasts perform [6]. Some parts of the forecasts were surprisingly good, others were surprisingly bad. What is clear is that it is not so easy to forecast how good the forecasts will be. Although the report greatly underestimated population growth in the United States, it was about right on the work week and productivity per hour. The resulting projection of 3% annual growth in GNP fell short of the actual growth rate of somewhat over 3·5% experienced in the quarter century that followed. The underestimation of materials requirements in 1975 which can be attributed to his under-estimate of economic growth is of the order of 12%. The motor vehicle population—an important determinant of materials requirements—was underestimated by 25%.

More serious was the Paley Commission's failure to foresee the rapid economic growth in the rest of the world. For continental Western Europe, the growth rate turned out to be 5%. For the United Kingdom the Commission assumed a 2% annual rate of growth as compared with actual growth of 3½%. For Canada the projected rate of 3% compares with 5·3% actual; for Australia 4% as compared to 5% actual, for Japan 4·8% with over 10%.

Since the Paley Commission projected to 1975 it is now possible to compare the projections of materials requirements with actual requirements. Interestingly, despite this underestimation of economic growth in the United States and its serious underestimation of the number of vehicles and telephones, the Commission more often than not *overestimated* United States materials requirements, by as much as 100% for cobalt, about 50% for tin, antimony and fluorspar, by about one-third for iron ore and chromite, and by about 20% for lead and managanese. It underestimated petroleum and aluminium consumption by about one-third and natural gas and electric power consumption by as much as two-thirds. It was about right on copper, zinc steel, nickel, molybdenum, tungsten and sulphur. The Commission's prediction of a resurgence of coal mining to about 800 million tons may yet come true in the next few years.

Because of a large under-estimate of economic growth in the 'rest of the free world' materials requirements here will be about twice the Paley Commission projections for copper, aluminium and sulphur: about three times or more for steel, chromite, nickel, molybdenum, manganese and fluorspar and four times for petroleum. Here the projections turned out right only for lead, tin and antimony. Although world-wide demand for materials increased much more than the Paley Commission anticipated, their predicted general upward thrust of real costs did not materialise.

While the GNP estimate for the United States by the Paley Commission was undoubtedly an excellent forecast the breakdown into sectors and specific materials (the object of the exercise) was far less satisfactory. The circumstances surrounding the making of the Paley forecasts were similar to those of recent years. Economies were strongly influenced by political factors (especially the Korean War) and commodity prices fluctuated widely. But even more recent forecasts, such as those of 'Resources for the Future' of United States demand for the rest of the century for individual metals and other raw materials show that uncertainty in demand is a factor of two after one decade and a factor of four after three decades [15, 6].

Economic growth

As a discipline concerned with the distribution of scarce resources, economics has had some successes compared with other areas of social science, although its reputation as the dismal science extends into its achievements in forecasting too. Part of the relative success of economic theory must depend on the availability of the monetary unit to describe a wide range of differing phenomena.

Vol. 5
1977

Even so, despite a fair amount of agreement between economists about the major variables affecting resource distribution (but less on their relative importance), our ability to forecast economic factors still remains very short-term.

There are intrinsic limitations to accuracy which arise from indeterminacy of measurement, changing patterns of behaviour and so on. Data will always be out of date or of a 'surrogate' nature. According to Bray [2] the limitation of accuracy of gross domestic product is about 0·7% of the growth rate for the ensuing year. Some idea of the amount of uncertainty in short-term forecasts

TABLE 2. VARIATIONS BETWEEN FORECASTS OF ECONOMIC GROWTH IN 1976. REAL 1976 ECONOMIC GROWTH

Country	DRI	Chase econometrics	Economic models	Forex	Percent OECD
Britain	1·2%	0·9%	2·0%	0·9%	0·0%
Canada	4·9	5·2	NA	4·5	4·25
France	4·4	5·2	3·6	5·5	3·0
Germany	5·0	5·6	5·0	4·1	3·25
Italy	1·9	3·7	3·1	0·6	1·5
Japan	4·2	7·1	4·3	4·4	4·25

Source: Business Week, April 1976.

using the more advanced models is given by Table 2. This shows the range of forecasts given by five different forecasting agencies during 1976, (*Business Week*, 26 April, 1976, p. 88). Thus, attempts to forecast national economies collectively (as in the OECD models or those of Economic Models Ltd) must be treated with caution. The opinion of many economists is that such work is 'premature'. When the most experienced econometricians (The Wharton Business School) refuse to enter their model for the race we should take heed. "The more uncertain people are and the less they know, the more willing they are to trust a computer model. But how much do we know of the world now? It's too risky" (F. Gerrard Adams quoted in *Business Week*, 26 April, 1976).

Although for long term economic forecasts assumptions about social and political trends are central, in fact, in the forecasting of life styles or the changes of our political and social structures we have made comparatively little progress. For concepts of quality of life, new international economic orders and the forecasting of surprises there are many theories which are contradictory but often more or less equally plausible [8]. For example, the length and severity of the current recession may suggest that more than a cyclical downturn in the world economy has occurred. For some this is evidence of impending limits to growth; for others it indicates only the need for structural reorganisation of the world economic order. Furthermore, there is the problem that unlike physical theory, social theory has an added

dynamic in that the theory evolves with society, and a new theory may itself cause society to change in an unpredictable way.

The physical sciences

Even in the natural sciences our predictive ability is quite limited. In physics, although so called fundamental constants such as the velocity of light, Plank's constant, the mass of the electron and so on have been identified, and in principle can be built into theories explaining almost any physical phenomena, in practice we are often in the position of the alchemist when one is attempting to produce such items as better semi-conductors, better super-conductors etc. We have a good understanding of the basic phenomena but this does not enable us to forecast as precisely as we would wish without checking conclusions with time consuming and expensive experimentation. Thus, although we largely understand the underlying principles of super-conductors this is only a guide; to discover the most suitable composition we make many different samples and test their properties.

In fact it is in engineering that the greatest precision in forecasting has been achieved but even so all is not plain sailing. We have a high level of measurability, a high level of reproducibility and because models are built on the same theory as the physical entity (a bridge or a power station or an aircraft) we are constructing, not surprisingly there is a strong similarity between the behaviour of the model and the physical entity. Nevertheless, we are very often in the position of having to build an experimental model, i.e. a pilot plant of the final system we are trying to construct or describe. The reason for this is that in some cases it is cheaper to build a model, although we may know in detail the rules to apply in a calculation (i.e. the underlying physical principles) the calculations are so massive or lengthy that we cannot perform them efficiently. (This is often the case in meteorology or hydrology.)

Vol. 5
1977

Holistic models

The question we must ask ourselves, therefore, is whether the level of theory necessary for the making of long-term forecasts is now sufficient, or likely to be sufficient in the future, for us to take the view that in forecasting we should presume that we are striving for accurate forecasts (even if these are of the kind what would happen if such and such took place?) Naive integrated forecasting appears to provide no certain answers. Nor does 'holistic' forecasting.

Take an example, relevant to current attempts to build large all inclusive ('holistic') Limits to Growth-type models [16]. Economic forecast based on purely economic factors could have led to an apparently 'accurate' forecast. If one recognised, however, that this forecast depended on assumptions about technological and political factors in which there was a great deal of uncertainty and then included these when forecasting ahead, one would immediately find

that, although we now have a better understanding of the way the world worked, we would in fact have much less certainty about the precision of our forecasts. We may, therefore, question whether our search for systematic explanations (of growth) of more than temporary or heuristic value may itself be doomed? The answer to this question is not necessarily yes or no but it does lead us to ask the direction of much research into forecasting methods.

Whether a particular forecaster feels more at home with quantified methods rather than qualitative or even literary methods will depend on his predisposition and experience. However, mathematical modellers in particular have many lessons to learn from the experience of, for example, urban simulation modellers who in the past have attempted to simulate complex socio-physical phenomena and in the main have had relatively little success. These modellers have believed that they were building models which would be useful to policy makers and in fact found themselves sadly disappointed when policy makers ignored their efforts. In addition there is the experience of researchers on national plans, for instance the French and Soviet Plans, who at various times have attempted to make detailed long-term forecasts and have found that not only for reasons due to absence of theory they have had little success [5].

At this stage I should also point out that however sophisticated a forecasting method appears to be, underlying it are always a few simple devices amenable to assembly into reasoned argument. These devices are concepts of measurement, the idea that the world can be explained in terms of simplified structures, that causal mechanisms operate which have a certain continuity both in time and place. Any method is reasonable only to the extent that these assumptions are fulfilled and can be represented by the method, i.e. they are valid only to the extent that one has a theory.

Of course, methods may help us to develop the theory and, for example, help us to deal with the complexity brought about by the interdependence of cross-national and cross-sectoral factors. Nevertheless, complexity leads to many other problems in the testing of a theory. From a methodological viewpoint it is clear from comparative studies [1], that despite the continued and increasing sophistication of economic forecasting methods there are diminishing returns to complexity.

Data is often collected for different reasons over short time periods in any non-homogeneous way and it is exceedingly difficult to differentiate short-term fluctuations from long-term trends. Our ability to test long-term forecasts especially is very limited simply because they are long-term forecasts. Short-term economic models, for instance, are tested on their track record, i.e. on their proven forecasting ability. Clearly with long-term models we can hardly do the same thing in a useful way and consequently we will usually be left with a great deal of uncertainty in our forecasts.

One of the most important features of a forecast for it to be useful is clearly some knowledge of its reliability, i.e. what it says about uncertainty and the

future. Many people argue that the level of uncertainty in the world is rising, one reason often given being its increasing complexity. This seems to me to be an incredibly difficult thing to judge. Is uncertainty about the future more or less than in Malthus day? The question is, however, irrelevant as it stands. What matters is our ability to cope. Perhaps by attempting to work with a range of theories that are, as far as we know, more or less equally valid we can achieve some measure of uncertainty. This is considered next.

Scenarios

There is a growing tendency now amongst forecasters to develop a range of images (or scenarios) of the future which indicate the outcome of alternative trends or specific contingencies and to test the 'robustness' of plans against them. One of the best known attempts to use scenarios in long-term futures research was by Herman Kahn in the book *The Year* 2000 [14] which, while it has been widely criticised for its forecasts, nevertheless contains some useful ideas about forecasting.

Of course, scenarios as such have been employed by decision makers since time immemorial. The basic idea of scenario analysis is to provide a context for thinking about specific issues—national or sectoral. One accepts (for all the reasons given earlier) that precise forecasting is not possible even to the level of giving a set of reasonably confident contextual assumptions. The answer is simply to consider a whole range of assumptions and to see what satisfactory strategies can be evolved, with reference to these differing sets of assumptions. Since the choices are arbitrary within a wide range of uncertainty several scenarios must be examined. (In *The Year* 2000 study by Kahn different kinds of international political arrangements were examined, ranging from 'fast growing development oriented worlds' to less appealing 'destruction dominated worlds culminating in all out nuclear war'). The task of the planner then is to find strategies in his own sphere of influence which are, as far as possible, compatible with the full range of anticipated outcomes.

However, an adequate assessment of uncertainty cannot really come from sensitivity testing carried out on a model which itself is based on a single theory, although most forecasters are satisfied with this. In both the studies of Kahn and 'The Limits to Growth' team certain variables are changed, although the underlying theory used to interpret data is not. Kahn's view of the world, for example, is focused on certain 'multifold trends' which he traces back to the eleventh century. (Broadly speaking he follows what may be termed the neo-classical functionalist approach. The *Limits to Growth* [16] and *Mankind at the Turning Point* [17] take a Managerial-Keynesian view while a Latin American world futures study [13] has essentially a version of a Socialist world view). But, as pointed out earlier, there are many ideologically competing views of history, all equally plausible (in the sense that all have a significant following) and each can be used to explore alternative futures.

Vol. 5
1977

More recent attempts at scenario writing try to take variations in theory as well as economic and other data, in an attempt to assess more carefully the range of possibilities which ought to be considered in thinking about the long-term future. In a recent American study conducted for the Environmental Protection Agency [9] a wide range of US futures were considered in evaluating alternative strategies with regard to environmental issues. At the Science Policy Research Unit, too, a similar examination of long-term world resource and technology requirements takes, as a starting point for discussion, twelve very different global futures [10]. Substantial guidance on the question of a suitable range of alternatives to be considered comes from reviews of past forecasts (such as those discussed earlier). Improvements in data collection and forecasting techniques can be taken into account also.

Scenario analysis is clearly an attempt to accept the fact of inevitable uncertainty in forecasting. It is also a recognition of the fact that even though scenarios may not always be much better than computer models they are also no worse, so nothing extra is gained from building a model. Further, they are less prone to spurious certainty which has the danger or restricting thinking thus causing alternatives to be overlooked. But the present retreat by long-term forecasters to scenario methods should not be seen as an act of despair; although the long-term future is impossible to forecast it need not be ill-considered. Neither does the claim that scenarios are often the most appropriate form of analysis mean that we should abandon attempts to build sophisticated models using, for example, complex computer methods. It does argue, however, that we should be very careful to identify the purpose of our work, in particular, we should not pretend that a model which in the theoretical sense must be viewed as highly tentative and experimental is sufficient to recommend dramatic policy changes. What models can help us do is to refine and compare particular parts of a given scenario or scenarios and their building can often help us to clarify ideas.

USE AND MISUSE OF FORECASTS

It is quite clear that the original *Limits to Growth* authors, in their well known 'prophecy of doom', did not appreciate this fact too well in making their 'doomsday' forecasts. It is also clear that their sponsors, the Club of Rome, did. To quote the chief executive of the 'Club', Aurelio Peccei, "what we needed was a stronger tool of communication to move men on this planet out of their ingrained habits" [11].

This then raises the question of under what circumstances does a forecasting method increase the theoretical or the social legitimacy of the forecast? In the past forecasts have been right for the wrong reasons, although often it is only the successful forecasters who are remembered. A forecast of dubious academic

quality may have an impact on a social process which may be self-fulfilling or self-defeating. Thus, very sophisticated forecasts may give much more credibility to a forecast than is really warranted. Even the most sophisticated audience can be misled, and to judge by the response to *The Limits to Growth*, they were. Too much certainty (especially if it is not justified) is at least as dangerous as too little! It is arguably impossible to keep propaganda and advocacy out of forecasting so attempts to separate planning and forecasting functions may be artificial. But for whatever reason a forecast is being used, the user at least must have some idea of its credibility.

Nor should this be taken to imply that all forecasts are intended to be misleading. Some are, but most are intended simply to lead. Ash and Smyth, for example, consider the possibility that published Government forecasts should be adjusted in order to encourage industry to act in a certain way, rather than simply to be a record of Government expectations [1]. Forecasts may also turn out to be misleading because of the terms of reference of the forecast. The notorious CEGB forecasts which systematically overestimated required generating capacity were partly the result of an expensive form of insurance to ensure that even peak demand would always be met and partly the result of an over-enthusiastic lobby. Given the need to play safe, overcapacity was the unavoidable result.

Vol. 5
1977

CUSHIONING UNCERTAINTY

There are, of course, many implications of uncertainty in forecasting and the approaches for dealing with uncertainty. From the point of view of the decision maker some essential lessons have been partially learned. That is to leave options open, not to commit resources until it turns out to be absolutely necessary to do so. Of course this is not always as easy as it sounds, many technologies and research programmes have very long lead times, so we must do the best we can to defend against uncertainty in the future. If a number of futures seem more or less plausible then we must be prepared to make a certain amount of investment to cover contingencies. One can take out 'insurance policies', here I am not simply talking about risk spreading through the insurance 'market'. For example, research and development has long and unforeseeable time periods for realisation. Thus, in setting up a programme to introduce a new technology we tend to concentrate initial investment on those areas which seem most risky and unpredictable before committing resources to those aspects of the project which are more easily forecast in terms of their cost and period for completion. Often one would here be engaged in parallel research projects until a particular line of development had proved its potential. To take another example from social theory, there are many different ideas about the future of education, and school design can account for this, i.e. classrooms can be built in a sufficiently

flexible form to permit, over the 20 or 30 years, the possibility of very different patterns of education. Here one is simply maintaining flexibility.

Essentially what these approaches consist of are ways of avoiding or mitigating the effect of uncertainty in forecasting. However, I should say immediately that one should not be too sanguine about any of them. For example, a reasonable observation might be that since a fair amount of uncertainty in forecasting comes from uncertainty about what other parties who are affected by and who can affect the forecast will do, the best procedure is to engage in mutual planning exercises. The OPEC situation is an interesting variant in this. If Western Governments forecast the collapse of OPEC energy saving will not occur and OPEC will continue whilst if they forecast its survival then it is likely to fail. Although in a bargaining situation it is often in one's own interest to conceal or distort information it is, nevertheless, possible to design so called 'positive sum' games which are to the benefit of all parties and make more efficient use of available resources. There are some obvious trade-offs to be made here. Some insurance strategies are merely protective, others buy protection and also contribute to overall welfare.

The question of whether Europe and other industrial countries should defend against possible producer cartels in major raw materials is an obvious example. When the current recession is over the emergence of OPEC type cartels are a possibility. "Insurance policies" can take the form of defensive strategies—stockpiling, R&D into substitutes, etc.) or into accomodative bilateral and multilateral agreements. Both strategies are likely to be expensive and both are vulnerable to economic or political circumstances but the latter can have a 'positive sum'.

One also takes lessons from the natural and the engineering sciences, where experiment is recognised as an essential counterpart to prediction. Indeed most legislation is, in practise, 'experimental' as it undergoes continuous modification. We should observe that social experiments as such are not intended, nor need, to have any sinsister connotations. They do, of course introduce matters of great public concern and as yet there is no satisfactory institutional structure for involving the public in such discussions, although there are useful experiments in this direction in the United States and Western Europe.

Obviously it is not possible to cover all contingencies satisfactorily either through design or through any form of insurance. The insurance market, for example, permits us to spread risks (especially those which can be assessed statistically) which are not easily avoided through reallocation of resources by government or industry. But it is clearly too expensive or too inhibiting of change to ensure *all* risks or keep *all* options open. Ultimately, we are left with a certain amount of uncertainty. This is a 'fact of life'. Nevertheless, to be aware of the dangers (and the symptoms) and to work towards their elimination is clearly a desirable objective and one which can be assisted through careful attempts at

forecasting. We are in a position to discuss alternative futures, mechanisms and the associated uncertainties but no magic methods can be expected to overcome the inherent difficulties of forecasting.

REFERENCES

1. ASH JC and SMYTH DJ (1974) *Forecasting the UK Economy*. Saxon House, London.
2. BRAY J (1975) Evidence given to House of Commons Sub-Committee. HMSO.
3. BUXTON M and CRAVEN E (1976) Demographic change and social policy: the uncertain future. Centre of Studies in Social Policy, London.
4. COLE HSD, FREEMAN C, JAHODA M and PAVITT K (1973) *Thinking About the Future*. Chatto & Windus, London.
5. COLE HSD, CLARK JC and ARCHER M (1975) Long-term forecasting in Western Europe. *Futures* 7(6), 485–514.
6. COSMAN CM (1974) Some thoughts on long-term supply-demand forecasting 22 years after. First World Symposium, Paris (mimeo).
7. COX P and WILSON G (1975) Problems of population modelling. Government Actuary's Department, London.
8. ENCEL S., MARSTRAND P and PAGE W (1975) *The Art of Anticipation*. Martin Robertson, London.
9. ENVIRONMENTAL PROTECTION Agency (1975) Alternative futures for environmental policy planning: 1975–2000. Office of Pesticide Programs, Washington, D.C.
10. FREEMAN C and JAHODA M (1977) *World Futures: The Great Debate*. Martin Robertson, London.
11. GILLETTE R (1972) Hard sell for a computer view of doomsday. *Science* **175**, 1088–1092.
12. GRIBBIN JR (1975) Climate, the Earth's Rotation and Solar Variations in *Growth Rhythms and the History of the Earth's Rotation* (ROSENBERG GD and RUNCORN SK, Eds.) John Wiley, Chichester.
13. HERRERA A (1976) *Catastrophe or New Society*. IDRC, Ottawa.
14. KAHN H and WEINER AJ (1967) *The Year* 2000. Collier-Macmillan, London.
15. LANDSBERG H, FISCHMAN L and FISHER J (1962) *Resources in America's Future*. Resources for the Future Inc., Johns Hopkins, Washington, DC.
16. MEADOWS D, MEADOWS D L, RANDERS R and BEHRENS W (1972) *Limits to Growth*. Universe Books, Massachusetts.
17. MESAROVIC M and PESTEL E (1974) *Mankind at the Turning Point*. Dutton-Readers Digest Press, New York.
18. US PRESIDENT'S ADVISORY COMMITTEE ON MATERIALS POLICY (1952) *Resources for Freedom*, (Paley Commission), U.S. Government Printing Office, Washington, D.C.

ADDRESS FOR CORRESPONDENCE *Dr HSD Cole, The University of Sussex, Science Policy Research Unit, Mantell Building, Falmer, Brighton, Sussex BN1 9RF, UK.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

Divisionalization and Transfer Pricing: A Review[1]

RB FLAVELL

Imperial College, London

(*Received March 1977; in revised form April 1977*)

This paper reviews the various transfer-pricing systems that have been commonly suggested for divisionalized companies. The properties of an ideal system are discussed and the advantages and disadvantages of each practical system are examined in relation to these properties. The fiscal restrictions that act on multinational companies are also considered. Finally, the contribution of operational research in establishing transfer pricing systems is discussed and criticized. Some very recent work is briefly looked at in the hope that this will form a basis for a more substantial contribution in the future.

INTRODUCTION

A TRANSFER price, *per se*, is a simple concept; it is the price charged for a good or service between two points within an organization. These points may be distinguished by time, physical distance and/or responsibility. The most clearly conceived form of transfer pricing is between two points separated by responsibility and this is the aspect that we shall concentrate on in this paper. However, before the mechanics of transfer pricing can be discussed in any detail, the nature of the responsibility separation must be examined.

Solomons [18, p. 3] draws a distinction between a 'decentralized' company and a 'divisionalized' one. Decentralization is, he claims, "the devolution of authority to make decisions" whereas divisionalization adds the "concept of delegated profit responsibility" to decentralization. Thus a division (or synonym) is a profit centre with responsibility not only for *how* its operations are performed but also, within limits, for *what* operations are to be carried out. The terms are of course inclusive; divisionalization must imply decentralization.

The major reason why companies are decentralized is *span of control*. This in fact forces decentralization on very large companies because of the bounds on the number and nature of problems that can be surveyed by any single manager and the number and qualities of people that he can manage. There is an optimum managerial size of any company, above which costs will begin to rise due to declining managerial efficiency and the need for additional coordination. A

[1] Presented at the OR Society Conference on Financial Management, Imperial College, April 1977.

second important reason, which is even more valid for divisionalized companies, is *motivation* and, linked in with that, management development. To quote Burlingame [2], divisionalization will create "a climate of individual responsibility, authority and dignity" that will provide a training ground for the development of the company's future senior management. Clearly, managing a division is in many ways similar to running an independent company and indeed in the very large companies, the divisions may be larger than many competing firms.

However there is one major disadvantage to divisionalization (and to a lesser degree, decentralization), namely *lack of central control* and hence company direction. This disadvantage is inevitable with the diffusion of central decision-making that is concomitant with delegated profit responsibility. The disadvantage is further compounded by the main criterion that is used to measure and monitor divisional performance—i.e. profit. There may be therefore a conflict between the divisional objective—to maximize its profit—and the company objective—to maximize its overall profit—and this conflict cannot be directly resolved by central authority without removing some of the delegated responsibility away from the division. The lack of central control is usually manifested by a lack of interdivisional cooperation; for example, a division may attempt to make an easy profit at the expense of a fellow division rather than go after a harder profit at the expense of a competitor.

In summary therefore, companies would regard divisionalization with mixed feelings. On the one hand there are the undeniable (but noneconomic) advantages of motivation and managerial development (coupled with the necessity for decentralization) and, on the other, there is the inevitable clash of (economic) interests between the central body and the divisional managers.

DIVISIONAL CONTROL

The usual measure of divisional performance is, as mentioned above, profit. However the dangers that this measure can bring must be minimized by any company wishing to divisionalize successfully. One way of course is to create entirely separate divisions that do not interact in any manner, e.g. in market, product, scarce resources or personnel. But such independence would imply that the divisions become so many unrelated companies and the central management merely operating an investment fund. To maintain the integrality of the corporate body therefore, the divisions must interact with each other, if only on a competitive level for the centrally controlled resources and facilities, but preferably at the interdivisional level as well.

Transfer pricing is a mechanism designed to provide some form of central control of the divisions without removing their delegated responsibilities. It operates by concentrating on the points of contact between the divisions themselves and between the divisions and the central management by determining

the prices at which goods and services are transferred. Having said this however, one comes to the crux of the problem: *how* are these prices determined? To illustrate the nature of the difficulty, suppose that all the prices are fixed by the central management; immediately a large portion of the *divisional* authority has been removed and the company is almost in a decentralized position with the accompanying motivational loss. Suppose on the other hand, that the prices are set by the divisions producing the goods and services; in this situation, these divisions could increase the prices of essential requirements of other divisions with the consequent transference of profits from the latter divisions to the former with the consequent transference of profits from the latter divisions to the former.

Clearly the determination of the transfer prices is of major importance to the company as a whole. The prices, however they are calculated, should in theory possess three properties:

1. Divisional autonomy must be maintained,
2. Divisions must act for the corporate good,
3. Divisional performance must be measurable.

These three properties will in practice almost certainly be in conflict and the next section of this paper will be devoted to a critical examination of various suggested methods with these properties in mind.

SUGGESTED TRANSFER PRICING SYSTEMS[2]

In this section we will assume as an example for discussion the existence of two divisions—the first (*A*) producing a semifinished product that is transferred to a second (*B*) who finishes the product and markets it. Other assumptions that are necessary will be made at the appropriate time.

Before turning to the individual transfer pricing systems themselves, it may be fruitful to consider the literature that has appeared on the subject. This literature broadly falls into two sets; the first deals primarily with management control and performance appraisal and generally appears from accounting and economic sources: the second discusses mathematical models of decentralized organizations from an operational research viewpoint. The accounting models in particular are pragmatic (see the first three systems below), concentrating on reality whereas the other models are more abstract and have not been widely implemented. A third set could be identified—relating to the fiscal requirements of transfer prices—and this will be discussed in more depth in the next section.

1. *Market price*

The price in the actual market for the semifinished product is widely regarded [9] as the most objective basis for transfer pricing. Its objectivity is that it

[2] This section is drawn from many sources; a major reference is Verlage[22].

introduces external competitive standards and measures divisional profits in terms of external opportunities (see also [17]).

If a perfectly competitive market does exist for the product, then, in theory, the pricing system will ensure an accurate measurement of divisional performance. Such a system would have to be modified in practice if either the amount of product being transferred is a relatively high proportion of *A*'s output or the system is to be operated for any length of time. This is because the transfers are intracompany, *A* is incurring minimal sales expenses and hence making supernormal profits that *B*, to some extent, is subsidising. It would not be long before *B* would ask for a discount on the market price in the light of this cost saving and of course the size of this discount is arbitrary.

In the absence of a perfectly competitive market, one could seriously question the objectivity of this system. For example, sales are not solely characterised by price, but also by delivery and terms of payment; the market price may be unrealistic—instances here would be dumping or the London Metal Exchange prices—due to large inter-divisional transactions and a small external market; and so on. Finally of course, no market may exist for this product and then the system is unworkable.

The system is widely used in practice by many companies, particularly those for whom a near-perfect market exists, e.g. basic raw materials. In the field of manufacturing companies, however, its use is dubious.

Vol. 5
1977

2. *Sales-minus*

The basis for this method is the sales price of the finished product to external parties. The method is particularly popular in practice when *B* is essentially a marketing division with little or no production facilities of its own.

There are two major disadvantages to this system of pricing; the first that the 'minus' on the sales price is arbitrary (although in theory based on some suitable overhead cost and profit margin) and hence open to negotiation, and second that it is incompatible with the concept of profit responsibility. For example, if the minus is based on actual (in an accounting sense) marketing cost, then increased turnover could lead to increased divisional profits via a very high marketing cost but not necessarily to increased corporate profits. On the other hand, if the minus is based on standard costs, there is an incentive for *B* to reduce the sales price and hence increase the turnover whilst keeping its own sales costs to budget. In this case *B* would increase its profits at the expense of *A*.

The system may be an acceptable method only if the sales price is *not* a decision instrument of the marketing division. But this has the inevitable effect of removing from that division its autonomy and profit responsibility and making it a pseudo profit centre [16, p. 832].

3. *Full cost*

Transfer pricing systems based on an allocation of costs can in practice be much more easily determined than those based on price. However, such systems do require the producing divisions to act essentially as cost centres which implies that divisionalization cannot be maintained.

The main advantages of a full cost (i.e. fixed and variable cost with or without a profit markup) system are that it is administratively cheap to run and easy to comprehend. The major disadvantage is that it does not normally operate towards the corporate good because the transfer price appears to *B* as a variable cost. This implies that *B* will not purchase the product (assuming of course that it has a choice) if it will make a loss on the product or if the product may be bought elsewhere more cheaply. But this decision rule neglects the fact that *B* will be contributing to the fixed cost of *A*—this cost will be incurred whether or not the product is transferred—and so the saving that *B* makes may not cover the loss that *A* will make. A second disadvantage is that the determination of full cost is more or less arbitrary; if actual cost is to be used, then this must either lead to delays or be applied retroactively, if standard cost then *A* will be open to the temptation to delay efficiency improvements until just after the (periodic) adoption of the standards. The above is valid without considering a profit markup: if a markup is applied then the arbitrary nature of this markup is an additional disadvantage.

This system will not, as we have seen, necessarily lead to divisions acting to the corporate good but if the amounts being transferred are relatively small then it is a very convenient system to operate. If the transferred amounts are significant, then the system is only really workable if the producing division becomes a cost centre and loses its profit responsibility.

4. *Marginal cost*

The previous three systems are basically accounting systems that have been widely applied in practice. A system based upon marginal cost comes primarily from enonomic sources [7] although it has been considered in some depth by accountants [12]. This system is more difficult to apply than the previous ones and so we will consider a very simple illustrative example; more complex examples are discussed in [12] but the essential mechanism remains the same.

For example, if *A* only sells to *B* and no market exists for the semifinished product, then the necessary calculations are as follows:

(i) determine the marginal cost curves for the final products of *A* and *B* (MC_A and MC_B respectively) and the marginal revenue curve for *B*(MR_B); note however that MC_B does *not* include the transfer price of the product but only the cost of *B*'s operations;

(ii) calculate the net marginal revenue schedule for *B* alone ($NMR_B \equiv MR_B - MC_B$);

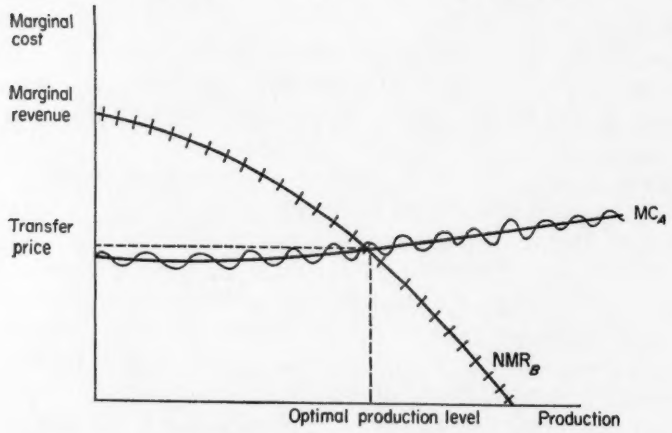


FIG. 1. *Calculation of transfer price.*

(iii) the optimal production for A is then given at the point where the marginal cost for A equals the net marginal revenue for B and the transfer price is given by the intercept on the other axis; see Fig. 1 for a graphical display of this method.

The advantage of this system is that it is theoretically optimal in the absence of a perfect or near-perfect competitive market for the semifinished product; optimal in this context implies that the divisions will always work to optimise the corporate good. There are, however, a number of disadvantages to such a system, some of which will be discussed here and others in the succeeding sections. The major disadvantage is that the transfer prices cannot be determined in isolation by A, nor would they emerge from a process of negotiation between the divisions without the marginal schedules being revealed. The only practical course of action would be to control the system from the centre but in this situation, are transfer prices at all necessary? If divisions have to surrender a considerable degree of their autonomy, why should the centre not dictate the volume of each product to be transferred? This would go to the heart of the problem! One answer to these questions might be that dictation of prices is not seen as such a serious infringement of divisional autonomy as the dictation of production policies, but this answer is not convincing [18, pp. 193–194].

A second disadvantage is that the system implies that all the units of the transferred product are charged at the marginal cost of the final unit. If the marginal cost increases with production, then this pricing will provide A with some contribution to its fixed costs and profit—which ultimately may or may not result in a net profit for A. If the marginal cost is constant (or decreasing), then clearly A will not recover its fixed costs (or not even cover its variable costs) B, on the other hand, because it may be gaining the product at less than full cost, could have a large cushion for its operations. This problem is of course

accentuated if there are several divisions in sequence, with the final one accruing much of the profits. Various methods have been suggested to overcome this disadvantage: e.g.

(i) equal profit sharing, i.e. if *A* produces an item for £5 and *B* sells it for £10, the profit would be divided into £2·1 for *A* and £2·9 for *B* (both of these figures give about a 30% return on the price at which it leaves the respective division);

(ii) contribution systems, i.e. *A* also charges *B* a periodic lump-sum payment to contribute to its fixed costs and profits;

(iii) a markup on the marginal cost, this markup again contributing to the fixed costs and profits of *A*.

Of course the disadvantage of all of these methods is that they are entirely arbitrary and may not add to the theoretical success of the system.

A final disadvantage is the problem, in practice, of calculating the marginal schedules, particular from a standard costing system based on constant variable costs per unit. It is possible to base a system on variable, rather than marginal, cost and use an adjustment such as (ii) above, but such a system would still suffer from the other disadvantages discussed.

5. *Shadow prices*

A further disadvantage to the previous system is that when dependencies exist between divisions or when there are active restrictions on divisional resources, additional co-ordination between the divisions is necessary for an optimal system to be implemented. To discuss this co-ordination, we must now move from the accounting and economic sources to the operational research literature; for a comprehensive review of the OR view of divisionalization, see [3].

In theory it would always be possible to construct a mathematical model of the company as a whole that would dictate the operations of each division so as to optimize the corporate good. This optimization could be carried out under the conditions, amongst others, that the dependencies and restrictions mentioned above would be satisfied. As a by-product of the optimization, shadow (or marginal or dual) prices for the transferred products would also result, and it has been argued [5, p. 184] that these shadow prices could form the basis for a system of transfer prices.

The shadow price system is of course a generalization of the marginal costing system, but with added conditions, and as such suffers from the major disadvantage discussed above. Another major problem is simply that the construction and solution of a sufficiently realistic model would be an enormous (if not impossible) task. However, see [4] for a much fuller discussion of the theoretical and practical problems of implementing such optimization models.

6. *Negotiation*

Finally a transfer pricing system based on negotiation is included in the discussion for completeness. Although it may be argued that the use of negotiation encourages competitive prices, it is likely that in practice, the most skilled negotiators will obtain the best prices, irrespective of the circumstances. Hence this system, albeit encouraging divisional autonomy, would certainly not assure that the divisions will act for the corporate good.

FISCAL RESTRICTIONS ON TRANSFER PRICING SYSTEMS

In discussing the restrictions that the fiscal authorities have laid down on transfer pricing systems, it is necessary to distinguish between transfers that take place under a single tax authority and transfers that take place across national (or in some situations, state) boundaries. For transactions that take place within a single country, the authorities have little concern as to the pricing system employed by a company (however, see the discussion on the US tax code below). This is because, assuming that the company is a single legal entity, the profits and losses of the divisions may be netted out and the overall company taxable profit computed. There are various schemes that may be adopted for reallocating the corporate tax bill amongst the divisions (see, for example, [18, pp. 116–122]): a common practice is an allocation on the basis of each division's taxable income. This allows the divisions to take advantage of any special tax treatments they may be able to obtain and it also credits loss-making divisions with tax allowances.

Vol. 5
1977

When one is considering transactions that occur across national boundaries, then the fiscal authorities take a far greater interest. This of course is not unexpected; every government considers itself not only to have the right to levy taxes on every business operating within its boundary but also in some cases to be able to claim additional rights, e.g. in certain circumstances the US Revenue Act (1962) taxed the profits of overseas divisions before the profits were repatriated as dividends. There are two main areas of interdependence between the transfer pricing system adopted and the fiscal authorities, i.e. customs valuation for import duty and the calculation of profit for taxation purposes. A third area is sometimes relevant when the eventual sales price to a third party is partially determined by the transfer price: examples here would be the roles played by the Monopolies Commission in the Hoffman La Roche case [10, p. 45] and the 'cadre de prix' in guarding the price of pharmaceutical products in France.

Before considering the above two areas in some detail, it may be useful to state probably the main requirement of most fiscal authorities—that a transfer price should be equivalent to an existing arms-length price, i.e. arms-length

transactions are preferred. If one considers the various transfer pricing systems described in the previous section, only the first three come anywhere near to meeting that preference.

Although transfer prices are not identical to customs valuation of the goods, it is customary to equate the two if possible to avoid delays at the port and to act as an acceptable basis for calculating the selling price to third parties. Hence, in considering any particular transfer pricing system, it is necessary to examine the customs regulations of the importing country. A great number of countries, including all the EEC countries but excluding the US, have accepted the Convention on the Valuation of Goods for Customs Purposes (Brussels, 1950), together with various recommendations made by the Customs Co-operation Council. These regulations laid down that the normal transfer price should be the market price for the goods as arrived at by an independent buyer and seller. This definition was enlarged and made more practical in 1968 [6] where it was recognised that the conditions of sales were very important and corrections could be applied to the market price before the import duty is calculated. If, however, a market price for the goods does not exist, and no comparable goods can be found, then a sales-minus approach should be adopted (the market price of the final product less the expenses incurred in the importing country and less a reasonable profit margin for the importer) to arrive at a valuation. Unfortunately, when an obvious market price is not available, the process often degenerates into bargaining between the importer and the customs authorities with no real reference to the value of the goods. A recent report [21] in the US recommended the adoption of a similar definition of value to that of the Brussels Convention, but at the present time, the basis for valuation is generally the fob price in the exporting port.

Turning our attention now to the problem of profit taxation, we see that transfer pricing is only one instrument (partially under the control of the company) by which the profits of the company may be moved around between the divisions. The entire problem, that of profit allocation so as to reduce the overall tax burden on the company, is a massive one involving the use of tax havens, bilateral tax treaties, central overhead charges, national differences in corporation tax rates, import duties and so on. (For more detailed discussions of these general topics, see [11 and 14].) As described in the above paragraphs, for items that are for sale on the open market, the transfer price is closely regulated by the customs authorities: it is in the pricing of the non-standard goods that the company may have considerable flexibility to move monies about. However, most countries have recognised the fact that it is profit allocation that is the real problem of the multinational company and have concentrated on legislation to prevent unacceptable (i.e. small) profits being reported. An important exception to this generalisation is the US which has attempted to prescribe acceptable transfer pricing mechanisms, again based on the arms-length principle [20]. This legislation, which is very detailed and contains

numerous examples, applies not only to companies with overseas connections (i.e. divisions or foreign parent) but also to internal transactions within the US. There is also a fall-back method: if none of the three acceptable pricing systems are applicable, then the Secretary of State has wide-ranging powers to calculate the allocation of profit, e.g. the total allocation of a foreign division's profits back to the parent on the grounds that the division is merely a 'sham'!

The UK legislation [19] is a complete contrast to that of the US; no operational definition of arms-length is given and the Section is simply a basis for profit re-allocation.

TRANSFER PRICING AND OPERATIONAL RESEARCH

We have so far discussed the advantages and disadvantages of various transfer pricing systems with reference to three ideal properties that they should possess, as well as outlining the legal framework within which multinational companies are able to fix the transfer prices. The general implications of this discussion were firstly that only three systems are widely accepted by the fiscal authorities—namely market price, sales-minus and full cost, in descending order of preference —and secondly that, in the absence of a perfect market for the transferred good, these systems will not ensure that the divisions act for the corporate good nor that the performance of each division is a true reflection of its overall contribution to the corporate profits. The approaches of marginal costing and shadow prices will ensure that the second implication will be satisfied but only at the expense of the divisions' autonomy.

Vol. 5
1977

There is obviously conflict between the three ideal properties and it has been suggested that the establishment of more than one transfer pricing system operating in parallel might be a suitable resolution. For example, Horngren [8, p. 540] suggests three systems—one for motivation, one for evaluation and perhaps one for dealing with third parties (the external prices)—but he does not consider the necessary administrative costs that would be involved in setting up these parallel systems. Nieckels [10, p. 156] also expresses his belief that such parallel systems would be cost-effective but does not verify his statement. In summary, therefore, this approach has obvious theoretical advantages but without considerable investigation its practicality is very questionable.

It is probably painfully evident that the three desirable properties are in conflict and will always remain so. The actual system adopted by any company must be a reflection of the overall managerial philosophy operating within that company. It may be therefore useful to examine the organizational structures of hypothetical divisionalized companies. Rutenberg [13] identifies four archetypes which he characterizes by the relations between the central management and the divisions. Building upon these archetypes, we have attempted to identify the effects on the three properties of the transfer pricing system that would be

TABLE 1. ARCHETYPES OF DIVISIONALIZED COMPANIES

Archetype	Organizational characteristic	Transfer pricing policy	Divisional autonomy	Corporate good	Divisional profit responsibility
1	No intervention by centre	Interdivisional bargaining	Complete	Ignored	Total
2	Indirect intervention only by centre	No imposed system	Considerable	Manipulated by centre	Very high
3	Direct intervention only by centre	Dictated by centre	Very little	Major objective	Functional responsibility only
4	Direct and indirect intervention by centre	Set by centre after negotiations with divisions	Collaborative	Major objective	Partial

implied: see Table 1. In the first two archetypes, there is no company-wide pricing system; the divisions are free to set their prices according to the local conditions of production and the results of negotiation with the other divisions. The centre plays a very subdued role except in archetype 2 where it attempts to influence the divisions by subtle manipulation. In both archetypes divisional autonomy is maintained to the detriment of the corporate good. The third archetype is dictatorial in nature, ignoring local conditions and where the divisions possess virtually no autonomy but are acting essentially as functional cost centres. Only in the fourth archetype is there any interchange between the centre and the divisions and it is interesting to note that Rutenberg comments "there are not as yet any examples of . . . corporations illustrating Archetype 4" [13, p. B347]; we doubt that the situation has greatly altered in the intervening eight years. Yet the classical decentralised models of OR have always concentrated on this last archetype, building mathematical theory upon mathematical theory and paying very little attention to reality. Atkins [1] arrives at a very similar conclusion and one could also draw a parallel here with the comments of Sandberg [15] about the organisational structure implied in most OR models.

So does OR have anything to offer in this area of study? Firstly, in the writer's opinion, the decentralized models do have a role to play, but on a *strategic* level and not an operational one. When a strong central management exists, as in archetype 3 and to a lesser extent in 2, then such models might provide useful indicators that would be fed into the general managerial decision-making process along with information gathered from other sources. The usefulness of these models is that they do provide a mechanism for examining the relevance of the interdependence of the divisions vis-à-vis the centre and each under other different corporate strategies. But the models are strictly limited as far as examining different transfer pricing strategies because the prices are not constrained explicitly as decision variables but implicitly as dual prices which in turn implies that neither the customs regulations nor the tax restrictions can be adequately handled. Rutenberg [14] constructed a simple network model of a multinational corporation, with transfer prices, managerial fees, short interdivisional loans and dividends explicitly considered together with certain fiscal restrictions, as a mechanism for deciding how the corporation's liquid assets could be moved in some optimal manner. The model is again a strategic one and Rutenberg suggests "one main use should be in reacting to anticipated devaluations, currency controls and other such risks and opportunities".

Another study that explicitly includes transfer prices as decision variables is that of Nieckels [10]. In this study, a non-linear programming model of a multinational corporation is constructed and an heuristic algorithm for solving the model is presented. But probably the main contribution of the study is that the model is then applied to a real company. Nieckels concludes by considering

Vol. 5
1977

different possible users of such a model, e.g. central management, divisions, governments and trade unions and discusses (as an example) how a government could investigate the effects of proposed legislation.

If operational research is to make a significant contribution to the problem of setting transfer prices and indeed to the wider problem of profit re-allocation within a divisionalised model, then it must address itself to the realities of the situation. There exists a vast reservoir of problems that could and should be investigated in conjunction with other specialists, e.g. what taxation strategies should a Third-world country adopt in the face of the multinationals, how should such companies take uncertainties of different types into consideration when planning, is there scope for the international taxation of multinationals? Such problems can only be examined by understanding the operations of the multinational companies. Operational researchers would approach such an understanding through the development of a model and at the present time there are (to our knowledge) no realistic models in existence. Nieckels' work [10] is only a first step; let us hope it is the beginning of a trend.

REFERENCES

1. Atkins D (1974) Managerial decentralization and decomposition in mathematical programming. *Opl. Res. Q.* **25**(4), 615–624.
2. Burlingame JF (1961) Information technology and decentralization. *Harv. Bus. Rev.* **39**(6), 121–126.
3. Burton RM and Obel B (1977) The multilevel approach to organizational issues of the firm: a critical review, *Omega* **5**, 395–414.
4. Dopuch N and Drake DF (1964) Accounting implications of a mathematical programming approach to the transfer price problem. *J. Acc. Res.* **2**(1), 10–24.
5. Dorfman R, Samuelson PA and Solow RM (1958) *Linear Programming and Economic Analysis*. McGraw Hill, New York.
6. European Economic Community Regulation (1968) No. 803/68, 27 June.
7. Hirschleifer J (1957) Economics of the divisionalised firm. *J. Bus.* **30**, 96–108.
8. Horngren CT (1975) *Accounting for Management Control* 3rd Ed. Prentice-Hall, New York.
9. National Industrial Conference Board Business Policy Study (1967), Interdivisional transfer-pricing, No. 122, L 68/27, New York.
10. Nieckels L (1976) *Transfer Pricing in Multinational Firms*. Almqvist & Wiksell International, Stockholm
11. Robbins SM and Stobaugh RB (1973) *Money in the Multinational Enterprise: A Study in Financial Policy*. Basic Books, New York.
12. Ronen J and McKinney III G (1970) Transfer pricing for divisional autonomy. *J. Acc. Res.* **8**(1), 99–112.
13. Rutenberg DP (1970) Organizational archetypes of a multinational company. *Mgmt. Sci.* **16**(6), B337–349.
14. Rutenberg DP (1970) Manoeuvring liquid assets in a multinational company: formulation and deterministic solution procedures. *Mgmt. Sci.* **16**(10), B671–684.
15. Sandberg Å. (1976) The role of rational analysis in decision-making: harmony or conflict perceptive. Invited paper at EURO II, Stockholm and published by EURO, Ed. by B. Bjorkland, Royal Institute of Technology, Stockholm.
16. Shillinglaw G (1967) *Cost Accounting: Analysis and Control*, Revised Ed. Irwin, Illinois.

17. SHILLINGLAW G and BURTON JC (1964) Materials management and the profit center concept. American Management Association Bulletin No. 4.
18. SOLOMONS D (1965) *Divisional Performance.* Irwin, Illinois.
19. U.K. INCOME AND CORPORATION TAXES ACT (1970) Section 485 and modified by the Finance Act (1975) which extended the definition of control.
20. U.S. INTERNAL REVENUE CODE (1954) Section 482 and also the Regulations issued by the Secretary of State, 16 April 1968.
21. U.S. TARIFF COMMISSION (1973) Customs valuation. US Government Printing Office, Washington, DC.
22. VERLAG HC (1975) *Transfer pricing for Multinational Enterprises*, Rotterdam UP.

ADDRESS FOR CORRESPONDENCE: *Dr RB Flavell, Department of Management Science, Imperial College, Exhibition Road, London SW7 2BX.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

The Chief Executive and his Information System

JC HIGGINS

R FINN

University of Bradford

(*Received March 1977*)

The authors examine the attitudes of British chief executives towards their information systems using recent survey evidence. They discuss the chief executive's individual computation needs and the extent to which he satisfies these personally, his attitudes towards computer-based information and corporate models, and his general informational and analytical requirements. Of particular interest, is the relative contribution as judged by chief executives of intuitive judgement, computer-based analysis and manual analysis respectively to strategic decision-taking; comparative valuations of formal and informal information systems are also ascertained. On the evidence provided by the chief executives, future prospects for computer systems and modelling techniques at strategic level are generally sound although it is difficult to see them achieving the dominance at top managerial level once envisaged by the more optimistic and imaginative management scientists and cyberneticians.

INTRODUCTION

In a previous article [3] the authors examined managerial attitudes towards computers, and in particular computer models, within the context of planning and control. The authors' conclusions were based largely on the relevant management literature, much of which is American, and they detected certain attitudinal differences between managers in the two countries, for example, in their respective views of corporate models. Furthermore, the view of chief executives are not widely reported. In an attempt, therefore, to contribute to the removal of some of these lacunae, the authors launched a survey, the results of which have now been analysed, of the attitudes of British chief executives and their head planners respectively.

A total of 56 organisations participated, the sample being drawn randomly from 'The Times Thousand Companies', and their industrial backgrounds are shown in Table 1. Most of the information was elicited by postal questionnaires but was usefully supplemented by some interviews and open-ended correspondence. Here our prime interest is in the views expressed by the chief executives. In a previous article in this journal [4], we examined the current

status of planning models in UK companies using the same sample of organisations.

TABLE 1. INDUSTRIAL CLASSIFICATION OF SURVEY COMPANIES

Industry	Number of companies
Agricultural supplies	1
Banking and finance	3
Beers, spirits and tobacco	2
Building and construction	3
Ceramics	1
Chemicals	2
Cinemas, theatres and television	2
Clothing and footwear	1
Coal, nuclear energy and steel	1
Communications	1
Data processing and office equipment	3
Electrical and radio	2
Engineering and metal	5
Food, catering, etc.	6
Glass products	2
Household products	2
Industrial plant manufacture	1
Insurance	2
Motor	1
Newspapers and publishing	2
Packaging	3
Services	3
Shipping	1
Textiles	1
Transport	3
Other*	2
Total	56

* Two companies did not identify themselves of the nature of their industries.

Vol. 5
1977

THE COMPUTATION NEEDS OF THE CHIEF EXECUTIVE

We posed a series of questions to the chief executives on the theme of their personal computation patterns and needs. The results for what we called 'numerical calculations' are shown in Tables 2, 3, 4 and 5.

The results tabulated are broadly self-explanatory and do not require great elaboration but several features are worth commenting on:

TABLE 2. FREQUENCY OF CALCULATIONS

Calculations performed	%
Frequently	38
Infrequently	51
Never	11

Frequently was defined as at least once a week, infrequently as less than once a week.

TABLE 3. SOPHISTICATION OF CALCULATIONS

Level of sophistication	%
a Back of envelope	53
b Quite sophisticated	23
c Ranging between a and b	13
d No calculations	11

TABLE 4. METHODS OF CALCULATION

Method	% using method
Hand methods	43
Slide rule	32
Log tables	2
Electronic calculators	68
Computers	17
No calculations	11

TABLE 5. LENGTHS OF TIME OF CALCULATIONS

Length of time	%
Less than 5 min	57
5–15 min	30
Greater than 15 min	23

(i) Less than two-fifths performed calculations frequently.
(ii) Only one in nine never carried out any calculations, however simple.
(iii) Of those who did perform calculations, they split in the ratio of 3 to 2 between 'back of the envelope' and more complex.
(iv) Three quarters of those carrying out calculations used electronic calculators and one fifth computers.
(v) More than one method of calculation was quite common.

(vi) Although over one-half of the calculations took less than 5 min, almost a quarter exceeded 15 min; the latter suggest some degree of sophistication.

Our next sets of questions were concerned with computations and analyses performed for the chief executive and, in particular, those involving some type of computer system. In Table 6, we have summarised the main reasons, and the percentages of chief executives giving them, for the delegation of decision

TABLE 6. REASONS FOR DELEGATION OF DECISION ANALYSES

Reason for delegation	% Giving reason
Insufficient time	29
Staff groups or subordinates have the necessary specialist knowledge	88
Computer facilities not readily available	4
Length of time required in data collection	6
Not part of the senior executive's role	25

analyses to a subordinate or staff group. The results tend to reinforce the evidence from the literature [5, 6] referred to earlier regarding the chief executive's chronic lack of time and liability to frequent interruption and his perception of his role. It is notable that lack of computer facilities is rarely cited as a reason.

ATTITUDES TOWARDS COMPUTERS

Almost two-thirds of the chief executives received computer outputs as part of the analyses referred to in the previous section. The overall results are shown in Table 7.

TABLE 7. FREQUENCY OF RECEIPT OF COMPUTER OUTPUTS IN DECISION ANALYSES BY SUBORDINATES

Receiving outputs	%
Frequently (at least once a month)	58
Infrequently (less than once a month)	29
Never	13

Table 8 shows the frequency of receipt of routine computer-produced reports. The answers to a set of relatively straightforward questions on some main features of these reports are summarised in Tables, 9, 10 and 11 and should require no further amplification.

TABLE 8. FREQUENCY OF RECEIPT OF ROUTINE COMPUTER REPORTS

Receiving reports	%
Frequently (at least once a week)	38
Infrequently (less than once a week)	42
Never	20

TABLE 9. EASE OF COMPREHENSION OF ROUTINE COMPUTER REPORTS

Reply given	%
Very easy	11
Easy	68
Sometimes difficult	18
Difficult	3

TABLE 10. USEFULNESS OF ROUTINE COMPUTER REPORTS

Reply given	%
Very useful	55
Useful	42
Not very useful	3

TABLE 11. SATISFACTION WITH VOLUME AND FORMAT OF ROUTINE COMPUTER REPORTS

Satisfied with	Volume (%)	Format (%)
Yes	74	86
No	26	14

ATTITUDES TOWARDS CORPORATE MODELS

Corporate models have had a chequered history on both sides of the Atlantic and it is manifest that they have not lived up to the early aspirations and expectations of many corporate planners and management scientists. Nevertheless, although the more grandiose dreams of total company models operating on strictly cybernetic principles, or with the ability to optimise the use of

resources across the whole entity, have now been put aside, there is general agreement amongst planners, management scientists and many accountants that a sensibly-designed corporate model is an invaluable top management tool. What though do top managers, in this case chief executives themselves, think of corporate models? In our sample, we found that almost three-quarters had some experience of the use of corporate models. The responses of all the chief executives to a question on degree of usefulness for strategic decision-taking are shown in Figure 1. It will be observed that over two-thirds judges corporate models to be 'useful' or 'very useful'; none stated they were 'no use at all'. It will also be noted that familiarity bred a higher rating than non-familiarity: for instance, half of those who felt corporate models were 'not very useful' possessed no direct experience of them.

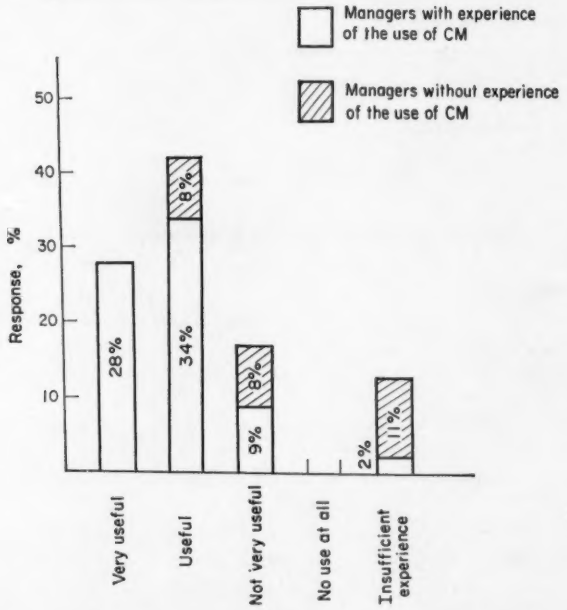


FIG. 1. *Usefulness of corporate models as aids to strategic decision taking.*

GENERAL INFORMATIONAL AND ANALYTICAL REQUIREMENTS

In the final section of our chief executive questionnaire, we broadened the scope to embrace more general views as to the nature of information and analysis regarded as important for top-level decision-taking and control. For example, respondents were asked to rate on a 10-point scale the relative import-

Vol. 5
1977

TABLE 12. RELATIVE IMPORTANCE OF FORMAL AND INFORMAL INFORMATION

Source	Mean score (out of 10)	Standard error
Formal management information	5.82	0.61
Informal ('soft') information	5.33	0.59

ance for their day-to-day information needs of formal information and informal information, much of which may be acquired by personal contact. As will be seen from Table 12, there is no significant difference between the values placed on the two types of information.

Now it is not infrequently stated by top managers that instinct and intuition are significantly more important in decision-taking at their level than formal analyses whether computer-based or not. We therefore asked the chief executives to rate, again on a 10-point scale, the importance of the following three factors in strategic decision-taking: intuitive judgement; computer-based analysis; manual analysis. Statistically, the differences between the pairs of scores, shown in Table 13, are not significant and we must conclude that the three factors are weighed roughly equally. Some results obtained by P. H. Grinyer and J. Wooller

TABLE 13. RELATIVE IMPORTANCE OF FACTORS INFLUENCING STRATEGIC DECISIONS

Factor	Mean score (out of 10)	Standard error
Intuitive judgement	5.75	0.54
Computer based analysis	5.33	0.53
Manual analysis	4.83	0.48

[2] make an interesting comparison: they asked corporate model builders in their survey to weight the relative importance of intuitive judgement, computer model reports and manual reports to top management in making decisions and found average scores of 43, 29 and 28 respectively. These are in the same rank order as our own figures for the perceptions of chief executives themselves but Grinyer and Wooller's planners allocate a significantly greater weight to top management's use of intuitive judgement. This may reflect an underestimation by planners of the extent to which top managers do use hard information and analysis; or it may be a result of some top managers either genuinely thinking they are more scientifically guided in their decision-taking than in fact they are or, perhaps, feeling that they should be and allowing this to influence their ratings consciously or unconsciously; and, of course, it may be that all these effects have contributed to the differences in relative scores noted.

We then looked at two possible hypotheses which might explain, at least in part, the relative lack of enthusiasm amongst some top managers for quantitative analysis: (i) Factors which may exert dominant influences on strategic decisions and their outcomes are sometimes difficult or almost impossible to quantify satisfactorily. (ii) An unstable business environment reduces confidence in detailed analysis.

The following statements were, therefore, posed and the chief executive asked to signify their degree of agreement:

A. In decision-making, non-quantifiable factors are frequently more important than quantifiable factors and detailed quantitative analysis is, therefore, not necessary.

B. In an uncertain economic world, one's own judgement is a more reliable guide to decision-making than detailed numerical analysis.

C. Computer based aids to top management decision-making are indispensable in the present business environment. The results are shown in Table 14.

TABLE 14. RESPONSES TO STATEMENTS A, B AND C

Statement	Agree strongly(%)	Agree with reservations (%)	Disagree with reservations (%)	Disagree strongly
A	0	24	62	14
B	7	34	55	4
C	16	55	29	0

The rejection by three-quarters of the sample statement A is consistent with the picture of a chief executive drawing on as much information as possible from whatever source, weighing each piece for importance and validity, and then making his final decision. There was no evidence in the answers to this statement of widespread doubts or scepticism as to the value of quantitative analysis.

However, the growing uncertainty of the business environment in the UK is reflected in the proportion, about two-fifths, agreeing that their judgement would be more reliable than detailed numerical analysis. GG Alpander's survey [1] of American executives showing that their use of management science techniques diminishes when they are posted abroad, seems to support this view.

Statement C was supported by about 70% of the sample suggesting that although many chief executives trust their own judgement more in the last resort, they nevertheless value the information and analysis a computer can provide.

Vol. 5
1977

FUTURE TRENDS

Most of the information collected was necessarily historical, albeit relatively recent, and current, and tends to reflect a situation in which the computer and modelling techniques have achieved satisfactory positions as aids to improved decision-taking, planning and control. We included questions as to future possibilities in two specific areas: real-time systems and corporate models. The respective responses are summarised in Table 15 and Table 16.

TABLE 15. FUTURE REAL-TIME APPLICATIONS

Application	Reaction to application (%)		
	Think they would use it	Do not think they would use it	Think they might use it
Try out strategic decisions on corporate model	46	32	22
Run programs of own specification	19	55	26
Make enquiries of records stored in the computer memory	47	32	21

TABLE 16. AUTHORITY FOR BUILDING CORPORATE MODEL

Response	Manager's experience of use of corporate models	
	Some (%)	None (%)
Yes	59	16
No	14	11

It will be seen that about two-thirds would, or might, use a real-time system to try out strategic decisions on a corporate model or to make enquiries of computer records. Fewer than a half would, or might, wish to run programs of their own specification. A few chief executives were already using such systems and the survey results clearly demonstrate considerable potential interest amongst the others. Given the evidence as to the nature of the chief executive's role, however, it is difficult to envisage him spending too much time at a terminal. One suspects that the attractions of direct computer communication might fade fairly quickly in many cases although, hopefully, the outputs from such systems will increasingly be made available to chief executives via subordinates.

The evidence for further development of corporate models is very positive, the figures in Table 16 representing percentage responses to the question

"Would you personally authorise the building of a corporate model, given the availability of appropriate resources, if your company did not possess one?"

In general, the gradualist approach to such developments adopted by British companies seems to be paying off and we would see a continuation of this trend rather than the emergence of any dramatic changes. There will, of course, be perturbations to the smooth course of innovation due, for example, to the impact of increasing environmental uncertainty on the attitudes towards quantitative techniques of some top managers or the economic constraints on expensive systems development. Nevertheless, the secular trend seems well-established in the UK.

REFERENCES

1. ALPANDER GG (1976) Use of quantitative methods in international operations by US overseas executives. *Mgmt Int. Rev.* **16**(1), 71–77.
2. GRINYER PH and WOOLLER J (1975) Computer models for corporate planning. *Long Range Planning* **8**(1), 14–25.
3. HIGGINS JC and FINN R (1976) Managerial attitudes towards computer models for planning and control. *Long Range Planning* **9**(6), 107–112.
4. HIGGINS JC and FINN R (1977) Planning models in the UK: a survey. *Omega* **5**(2), 133–147.
5. MINTZBERG H (1975) The manager's job: folklore and fact. *Harv. Bus. Rev.* **53**(4), 49–62.
6. STEWART R (1967) How managers spend their time. *Mgmt Today* (*June*) p. 92 *et seq.*

ADDRESS FOR CORRESPONDENCE: *Professor JC Higgins, Management Centre, University of Bradford, Emm Lane, Bradford, West Yorkshire BD9 4JL, UK.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

Criteria for the Selection of Reporting Method

JR MACE

University of Lancaster, UK

(*Received March* 1977; *in revised form April* 1977)

Identification of optimal reporting method by analysis requires that a criterion be available for selecting those 'best' practices which are to act as the norms for standardisation. 'Predictive ability' has been suggested as such a criterion. This article argues that 'predictive ability' should be rejected in favour of an overriding economic criterion for the design of reporting methods and information systems. An information system is only valuable as part of a wider system. This article proposes that the wider system in which we are interested must be 'more valuable' in contemplation of change than without it, if changes in information systems are to be justified.

INTRODUCTION

THERE is fairly general agreement that accounting standards are desirable in order to narrow the areas of difference and variety in accounting practice, and to reduce the ambiguity of accounting reports in their representation of the affairs of companies. The work of bodies such as The Accounting Standards Steering Committee in the UK and the Australian Accounting Standards Committee seems to be based on the belief that standardisation of practice will enhance the usefulness of accounting reports for transferring knowledge of company affairs from those who have it to those who want it. A preliminary step to the formulation of accounting standards must be an agreement as to the objectives which accounting reports are intended to serve. [8, 10, pp. 147–148]. Without such agreement there can be little hope that standardised reports will be more successful than non-standardised reports in generating the same responses in their users as would have been generated in them if they had experienced personally the events and situations to which the reports relate. To achieve successful standardisation, we require an agreed method for selecting those 'best' practices which are to act as the norms for standardisation. The purposes of this paper are to challenge recent claims that searches for 'predictive ability' in accounting data will identify optimal reporting practices [e.g. 4, 5, 12,

25, 26] and to put forward an alternative to predictive ability as the overriding criterion for the selection of a standardised reporting method.

The article will claim (i) that the predictive ability criterion is as yet insufficiently precise for practical application, (ii) that precise tests of predictive ability (when they are acceptably formulated) will be of extremely limited practical importance, (iii) that the predictive ability criterion is subservient to an economic criterion for the design of information systems, (iv) that proposals for making recommendations on accounting practice based upon testing the predictive ability of certain accounting numbers with respect to the information requirements of investors in ordinary shares are mistaken because (a) there is no consensus as to the informational needs of investors and (b) there is no agreement that the needs of any one group of users of published accounting information should be given preference over the need for accounts to provide information to decision takers generally.

THE PREDICTIVE ABILITY CRITERION

Vol. 5
1977

We begin by examining the literature concerned with predictive ability as a criterion for the selection of accounting method. 'Predictive ability' has been suggested as 'an overriding and justifiable criterion' in assessing the ability of various accounting models to meet the decision-taking objective of their users [15, p. 273]. The predictive ability criterion does not regard the accounting numbers as predictions in themselves; it says that one measurement system is to be preferred to another if it produces numbers which are more useful inputs to a forecasting process [15, p. 279, also 6 13, 25, Ch 4 and 5]. Special purpose researchers have proposed to apply this criterion to a number of accounting measurement methods to determine the best accounting method for predicting those variables of interest to the users of financial statements [2]. They propose that the specific needs of a number of separate user groups for each of their decisions will be examined in this way as a preliminary step to discovering whether any recommendation is possible concerning the content of accounts to be published for the facilitation of decision taking in general. They have begun their work by concentrating upon seeking an accounting formulation which will best predict the future events and sequences of events required as inputs for the portfolio analysis model.[1]

They claim that this choice is justified because the external financial reporting

[1] The search is claimed to be subject to general constraints concerning the cost of providing information, the confidentiality of information and a social welfare objective of Pareto optimality for the state following any changes which may be proposed. No discussion is given of the method of achieving the desired *general equilibrium* analysis in place of a *partial equilibrium* analysis, or of the way of identifying and judging the impact of accounting decisions upon the wealth of all members of society which is the pre-requisite for testing the Pareto optimality criterion.

of figures with predictive ability would further the objective that unbiased expectations should determine the market price of a company's shares.

The construction of a normative decision model for a particular type of decision by a decision taker is based upon an assumed and precisely described objective which enables the researcher to deduce an appropriate structure for the decision model.[2] He is then able to consider the several ways in which the information inputs may be provided to the decision model and to apply a chosen criterion for choosing between them. The process of working from decision model to information source may not be the 'natural' one for the development of the joint system of decision model and information source. It is perhaps more likely that decision-taking methods evolved to use *available* information and that actual decision methods have been based on observed characteristics of available information. A feedback process between the user of accounting information and the compiler of accounting reports would ensure that the usefulness of the information produced reflected a point of balance which had been reached in a bargaining process concerned with the needs of the user and of the compiler. A limit to the effectiveness of the feedback process may be summarised as 'the inability of the user, from his experience alone, to instruct even a willing accountant as to the information he needs *which an accounting system can produce*' [9, p. 162]. It may not be feasible for the accountant to produce the ideal information requirements for a particular decision model. The researchers propose to deal with this problem by finding the best available substitute for the ideal information requirements specified by *their* model rather than seeking to provide required information inputs for *existing* decision models in a more efficient way.

Any statement of the predictive ability of an accounting method must be qualified by a statement of the usual assumptions upon which statistical prediction processes are based. Any proposal for change based upon studies of predictive ability must be modified in relation to the *net value* the change will have to the economic system in which it is to be applied. To find an accounting method with predictive ability is a matter of disinterested reasearch but the advocacy of the method in the real world is different. It is a political act, since the problems of social choice are not amenable to analytical solution. A researcher who wishes to gain the support of practical men for the proposals he advocates must ensure that his research model has the characteristics of the practical man's perceptions of the real world [20]. Special purpose research based on the portfolio analysis model has several deficiencies in this respect.

An approach to a practical problem by using the normative portfolio analysis model is of doubtful validity (even within the general limitations of all special purpose research) since the measure of mean return and the measure of risk to which the portfolio analysis model should be applied have not been empirically

[2] Note the distinction between provision of information for a decision model and provision of information for a decision taker.

determined. In short, the theory is not sufficiently well defined for practical use, since it does not specify what observations are to be made in practice or what properties of the real world are to be measured. The truth of the portfolio analysis model is tautological. For any consistently defined use of the terms 'risk' and 'return' the model will produce a portfolio having an optimal combination of risk and return in relation to a utility function expressed in terms of risk and return. What has not been done is to specify what types of risk and return give utility to a *living* decision taker and to show how *those* types of risk and return may be measured or estimated in the real world.

Few organisations yet use net present value techniques for investment appraisal, despite the theoretical arguments in their favour (though possibly for valid reasons), so we may expect that special purpose researchers will find it difficult to convince practical men to produce accounting information for the portfolio decision model (if such could be found) rather than for general decision taking purposes in a way that attempts to satisfy the conflicting interests of individual users of accounting information.

The practical usefulness of the solution derived from treating a subsystem of the real world as a closed system is largely dependent upon the degree of independence actually present (but this is of course unknown) [21].

In the view of this author, the optimisation of the widest system in which the practical man shows an interest (and of which the system under examination is a part), is more likely to be achieved by recommendations derived from a general purpose approach than from a special purpose approach. A researcher who examines a possible structure for general purpose accounts is making a direct attack upon the problems of the real world. The general purpose approach promises more immediate practical results than the analysis of special purposes because it omits the research stage in which the results for independent user groups have to be reconciled with one another (perhaps by a form of analysis akin to game theory), and because it does not require the explicit consideration of all types of decisions before practical proposals may be made.

Recommendations derived from general purpose research are likely to be more relevant to the political process from which accounting standards emerge than those derived from special purpose research because the system which is implied as the framework for the research is one in which the interaction between competing uses and competing users of accounts is explicitly recognised *ab initio*. In other words the general purpose model does not involve as much abstraction from the real world as does the model of the special purpose researcher.

THE INVESTOR'S INFORMATION NEEDS AND THE PORTFOLIO ANALYSIS MODEL

Let us assume for the moment that there is an underlying probability distribution for an observable variable known as 'the financial performance of a company', and that the variable is of interest to investors. The purpose of discretionary managerial action is to control this variable, by taking decisions periodically and putting them into effect. The time interval between consecutive reviews is empirically determined by economic and social factors. The value of the performance variable will be an unknown function of (i) previous performance, (ii) the initial state of the company, (iii) current managerial decisions, and (iv) environmental and intra-organisational change during the period. The minimum object of performance prediction will be to create an estimate of the performance during the decision period and of the likelihood of error in the estimate. The planning process is not restricted by the length of the decision period and it is usual for plans to be prepared which extend beyond the end of the current decision period to a point known as 'the planning horizon'. A sensible planning horizon is the point beyond which the cumulative effect of environmental and intra-organisational change is expected to be so great that it would be uneconomic to prepare additional plans at the present time. The location of the planning horizon is therefore dependent upon the stability of the relationship between the performance of the company from one decision period to the next. It is limited by (i) the uncertainty of environmental and intra-organisational change and by (ii) the expected effectiveness of discretionary managerial action in providing control over performance. In other words, the planning horizon is influenced by beliefs about the statistical independence of performance from one period to another. An actual planning horizon for outline plans is rarely likely to exceed 5 or 10 years and detailed plans would normally be prepared for a much shorter period than this. The prediction of performance may extend beyond the decision horizon and will normally extend forwards to the planning horizon.

Consider the need of an investor for predictions. This need will depend upon his decision methods, his objectives and his views of the market in which he operates. The possible number of combinations of assumptions about these variables is at least as great as the number of investors. Any discussion requires that we concentrate on just a few of the possible combinations. If our discussion is to have empirical content, we should select a combination of assumptions which is widely accepted as realistic. We may assume, as a starting point, that (i) all companies have similar decision horizons and planning horizons and that their timing of decision taking is in phase. (ii) the investor has a statistically stationary two-parameter objective function and seeks to maximise the utility he obtains from the expected return and risk of his portfolio of investments. (iii) the investor is fully aware of his utility function. (iv) the decisions and invest-

ment actions of the investor are instantaneous and costless. (v) the investor views himself as a price taker, and (vi) there is no taxation. In order to apply the normative portfolio decision model, the investor must predict the relevant risk and return characteristics of each share which is perceived as a potential investment. For this purpose, he requires (i) to know the relationship between market price and expectations about the future performance of the company at the time of his decision, (ii) to predict the actual performance of the company during its decision period and the standard error of his estimate as a measure of its risk, (iii) to estimate the future relationship between market price and expectations about the future performance of the company as it will be at his decision horizon, when expectations have been revised on the basis of additional experience of the company's actual performance and (iv) to estimate the covariances between the returns of all of these securities in the period ending with his next decision. In this paper we shall restrict ourselves to a consideration of the information needs numbered (i), (ii) and (iii) above, which are required for the calculation of the expected return on a single security.

Information requirement (i) may be resolved by means of a valuation model such at that which defines value in terms of the discounting of expected cash flows. This valuation model enables an investor to 'explain' the present market price of a share or to arrive at a present value for the share on the basis of his own estimates of expected cash flows and a discount rate. The value so derived may be called 'the subjective value of the share', since it is the current value of the share from the investor's point of view in his current state of knowledge. Calculation of the subjective value of a share is without purpose unless the subjective value is believed by the investor to be a better estimate of the present value of the share than is the current imperfectly determined market price. Consider how the subjective value is used in taking investment decisions. The difference between the subjective value of an investment and its market value or the 'cost' of its component parts is known in the literature of business finance as the net present value of the investment. Given a suitable definition of the appropriate discount rates for use in the valuation model, the absence of a negative net present value is the criterion for acceptance of the investment.

The investor's estimate of the expected return on a share until his decision horizon (i.e. during his holding period) will take into account the changes in the market price which he expects will occur during the period. Implicitly, in his estimate of expected return, he is deciding how newly available information about the company's performance may have affected the market price of the share by the end of his holding period. An excess of subjective value over market value will be relevant to his expected return if some change in the proportionate excess is expected during the holding period. The expected return on an investment at the current market price will increase, for example, if the excess is expected to be reduced by movements in the market price during the period. Thus the return expected by the investor depends *inter alia* upon the

Vol. 5
1977

extent and direction of the difference between his subjective evaluation of the share and its market value, and upon his estimate of the extent and direction of the rate of change in this difference up to his decision horizon.[3]

This formulation of the investor's information needs differs from that which was presented by the special purpose researchers to whom we have already referred. They asserted that a shareholder's main interest was in the level of future dividends, they claimed that knowledge of future operating cash flows, together with a knowledge of past dividend policy would be likely to provide a good basis for the prediction of future dividends, and they proposed in effect that alternative accounting standards should be judged on the basis of their ability to predict operating cash flows.

The prediction of distributable operating cash flows must be combined with predictions of future dividend policy if future dividends are to be predicted. Future dividend policy is susceptible to government control which is only indirectly related to the economic situation of individual companies, and to radical amendment when changes of control occur or are threatened. The accuracy of prediction of future dividend policy in relation to distributable operating cash flows is of critical importance in determining future dividends and there is little to suggest than an elaborate prediction model based on information from accounting reports will improve *on the average* on results obtained from a naive extrapolation of some time series of dividends. In another article 'relative uncertainty' as a measure of risk in the context of portfolio investment has been presented as an additional object of prediction [2 pp. 99–100].

It was also suggested that control information was a necessary input to the decision model, though it was not conceded that the provision of optimal control information may be inconsistent with the provision from the same source of optimal information for the prediction of future operating cash flows and their relative uncertainty.[4] The researchers have not yet considered how future relative uncertainty may be predicted from accounting reports, though they suggest [7] that the construction of a model for predicting future distributable operating cash flows should involve a direct look at the process which generates cash flows.[5]

[3] These points are usually ignored in the exposition of the information requirements for the application of the portfolio analysis model in practice. See J. A. Arnold and A. Hope [2] for an example of this omission. They propose to follow Sharpe's simplification of the original Markowitz model and introduce the concept of an efficient market portfolio in order to avoid making estimates of the covariances between the returns on all securities perceived to be in the market, though they do not consider the practical problems of identifying an efficient market portfolio in the real world.

[4] More recent work has considered *a priori* arguments suggesting that economic income has the greatest predictive value and takes the view that an optimal accounting method would be the one which produced income numbers which have the most consistent relationship to measures of economic income.

[5] Lev also makes the point that the identification of the statistical process actually generating corporate earnings is crucial for the construction of prediction models [18, p. 125].

THE IMPRECISION OF 'PREDICTIVE ABILITY' TESTS

The testing of predictive ability must begin by specifying what event or events are to be predicted. We have argued that the determination of the information needs of an investor is a difficult task even when the literature contains a well-established normative decision model such as that for portfolio analysis.

As we have seen, the object of prediction requires careful specification. The sequence of operating cash flows which is required to predict dividends for use in the normative valuation model extends to the termination of the enterprise. Only simple models have yet been put forward to predict that far into the future. As investors in failing companies often find to their cost, simple models tend to predict to an indefinite future for a time series, rather than indicating *the timing of a definite end* to the series. It is difficult enough for businessmen to distinguish between temporarily bad times and an inefficient business, and it will be even more difficult for the investor to superimpose upon this prediction a prediction of when the businessman will perceive the truth and take appropriate action. We cannot wait long enough to test the efficiency of methods of predicting an indefinitely long time series, so any practical test is necessarily restricted to a test of the accuracy of prediction for a limited and arbitrary number of periods ahead of the time at which the prediction is made (though preferably for the period long enough to extend to the time horizon of the decision taker). It is usual for the accuracy of predictions to be considered in terms of lowest probability of error or lowest standard error of estimate in prediction of a single number (being a time series variable or a parameter of a distribution). Alternatively, we may evaluate accuracy in terms of the highest correlation between predicted numbers and actual numbers or in terms of the lowest average informational inaccuracy of forecast.[6] Such methods fail to take into account *the loss of value* associated with actions taken on the basis of predictions which are in error, and are therefore unsuitable guides to choice of a prediction method which must ultimately satisfy a value maximising criterion. Effective valuation of the merits of alternative prediction models must include a direct measurement of the value of actions and of the cost of errors. Absolute differences between a deterministic prediction and the actual outcome (or standard variance analysis of such differences) will not be appropriate measures of the accuracy of the prediction, nor will they be sufficient stimuli to signal the direction of modification in the parameters and variables of a real world prediction model for its improvement and adaptation as the environmental conditions change.

Consider the nature of the prediction process. Goldman refers to the opera-

[6] For an example of the use of the latter criterion, see [29].

Vol. 5
1977

tion of a prediction model as a method of analysing the flow of information from a source in order to predict what is going to happen on the basis of past experience. The flow of information is a discrete or continuous sequence of measurable events distributed in time and its analysis involves studying the historical sequence of signalled symbols in order to determine the statistical characteristics of its intersymbol influences. From the past and present values of a signal, we can then predict the future of the sequence up to the end of the observed range of intersymbol influence. The prediction will decrease in accuracy, the further we get into the future and only an average value may be predicted when we go further into the future than the range of intersymbol influence. The decoding of received symbols is an application of scientific method. The interpretation of signals requires memory and the meaning of *new* signals may not be communicated but must be learned from subsequent experience. No manipulation of received signals without further contact with their source will on the average increase the technical level of information they contain [19, p. 90]. The information *which is perceived* from a given set of signals may nevertheless be enhanced by mathematical manipulation of the signals received, since the new formulation of the same technical information may permit further interpretation of the signals to occur.

On this view of the information contained in signals, it is impossible for there to be any interpretation of publicly available signals which is not itself publicly available, except through some previously untried mathematical manipulation which acted as an original tool or technique of analysis. 'Interpretation' and 'prediction' are closely associated concepts. Interpretation of symbols and of symbol sequences consists only of prediction on the basis of history repeating itself, prediction by association of similar events, or prediction by analogy on the basis that a complex situation may be likened to a simpler one. In an 'efficient' share market, the prices of ordinary shares would fully reflect all available information at all decision points, and no subjective evaluations could be expected to provide reliable guidance for portfolio choice. In an efficient market, the investor must construct a valuation model (perhaps a dynamic model) which explains current market prices and predicts future market prices. It is argued that past variations from previously expected prices produced by the model would be the best available estimate of future variations from expected future prices, and hence the best estimate of the risk associated with the expected return from holding the share. The argument is attractive but it is false. It ignores the fact that the investor has no basis upon which to judge whether an observed difference between a predicted return and an observed return signals the risk of the share, or a fault in the existing valuation model. It also ignores the fact that the only measure of risk which is of interest to the user of the portfolio analysis model is a function of his expectation and may not be related to any measure of variability of *previous* outcome. Furthermore, as we have seen, a model for the estimation of the future price of a company's

shares relies upon untestable assumptions about the formation of cash flow and discount rate expectations from observable experience.

Lev writes that extensive empirical evidence upholds the theory that stock prices on large stock exchanges fully reflect all publicly available information and adjust rapidly in an unbiased manner to new signals. The major objective of an analyst's investigation of signals in such a market is to obtain information that is as yet unknown to the general public. Extensive use of any analytical method for uncovering inside information will soon result in market prices reflecting the value of information provided by the method and thereby gradually decrease its usefulness. Lev's conclusion is that originality and secrecy in developing analytical tools and alertness in using them are the major conditions for successful security analysis in efficient capital markets [18, pp. 222–223]. Any well-known or 'proved' analytical tool is probably useless by the time it is published since the prediction it permits is not statistically independent of the predicted event.

Simulated data may be the foundation for the evaluation of the predictive performance of a number of alternative prediction methods, but the 'best' method under the simulated conditions can only be the best method under real conditions if the simulated data is typical of real world data at the time of its use and if the method of expectation formulation used in the model is typical of the real world method. The limits to statistical analysis of signals are that any reasonably large sample should tend to be representative of the signal sequence as a whole (i.e. of its whole past and its whole future) and that the predictions are independent of the predicted events. If simulated data is typical only of a limited time period in the real world, the optimal prediction method in the simulated situation will only be optimal when similar data sequences are produced in the real world. There can be no assurance that past predictive performance will continue into the future if the accuracy of predictions by the model is sensitive in the real world to the time period involved, or to small changes in the definitions of the variables whose measurements are required by the model, or if the model fails to predict turning points in trends of events. It is familiar to statisticians that the normal methods of analysis of a time series permit little more than the identification of seasonal fluctuations from an established trend. They know that any predictive advantage one may expect to derive from distinguishing turning points and changes in trends from longer runs of statistics will be specious and spurious if those statistics (like data from past accounting reports) have been collected under widely varying conditions.

All prediction is based upon the persistence of observed phenomena but we may question the validity of this assumption where business data is concerned. The retention of historical records of past events (and in particular of past *decisions* and their succeeding events) provides some knowledge and experience in the control process. It also provides a basis for the prediction *ceteris paribus* of the events which will succeed the taking of similar decisions in the future. In

Vol. 5
1977

this case, the validity of the prediction depends upon the assumption that there has been no material change in exogenous variables and before any reliance is placed upon the prediction, efforts will be made to verify that such changes have not occurred. In the business situation, it may be less easy to enumerate all the relevant variables than it is in the case of the physical sciences. In practice, we minimise the likelihood that material changes have occurred in exogenous variables by concentrating upon the most recent historical experience as the basis for predictions and by discarding old data.

The accuracy of predictions made by such methods is limited as a result of those *essentially new* events which occur in the historical sequence as a direct result of human attempts at problem solving [20]. The optimal base period for any predictive model may only be found by trial and error. It extends from the previous relevant *new* exogenous event, to the present, though it may be shorter if the greater cost of predictions based upon the whole of such a period is not justified. The identification of the new exogenous events which bring about the end of the lifetime of such prediction models is an important aspect of their successful use. Historical sequences of accounting records are important for the prediction processes necessary for control because they specify the base point or starting point upon which speculation about future events must be built and from which any predicted change must take place. The numbers used in accounting reports may often be used as predictors of the opportunity costs which would be among the ideal inputs for many decision rules concerning the allocation of resources. The latest set of status records will be the best base point for predictions (independent of any part it may have in establishing or confirming a trend in a sequence of figures). The latest set of status records will be an adequate substitute for opportunity costs applying at the time decisions are put into effect *provided* current costs and prices are used in maintaining the status records.[7] For other objects of prediction, such as the identification of important exogenous events (see above) or the use of accounting records for control purposes, the optimal accounting method may involve *a priori* the use of some form of constant cost pricing.

The validity of an academic search for a carefully defined type of predictive ability has not been in question, but its practical value is questioned when an alternative use of research resources would be to try to solve the practical problems of implementing a socially accepted *a priori* conclusion about the desirability of current cost pricing.[8]

The objective of the special purpose researchers is to find an identifiable relationship between what is measured and reported, and their chosen object of

[7] For discussion on the concept of opportunity cost, see [1]. Note that recently recorded historical costs would only be a substitute for the *external* opportunity costs of resources and that internal opportunity costs will be relevant when resources are rationed and must be estimated in other ways.

[8] See for example, the proposals in 'The Corporate Report' [28, section 7].

prediction, and they have the hope that the identified relationship incorporates time lags in a way that will help in the prediction of turning points in the trends of economic events. It is natural to set up an ideal in which 'lead indicators' will act as warning signals of changes in trends but doubtful that such an indicator will retain predictive ability for long in a real world in which actions may be taken which prevent or alter the timing of the event which has been predicted. In other words, we must distinguish between predictions of two types; the first is concerned with events in the decision period and the second is concerned with events up to the planning horizon. A prediction of the first type is that for which the *ceteris paribus* assumption is appropriate. It cannot be tested in a world in which the decision horizon of economic actors is unknown and decision takers are believed to adapt their behaviour in accordance with newly available prediction techniques. A prediction of the second type, while testable in the real world, is unlikely to be derived from a model which may be continuously applied in the same form without change because the attempts of humans to solve their problems will create situations with significant features which are inherently new [20]. Predictions of the second type must take into account such things as the learning capabilities of economic actors and the implications of publication of predictions of the first type. The validity of prediction methods leading to predictions of the second type is inherently difficult to test since it is unlikely that adaptive behaviour induced by publication of such predictions would be neutral in its effect. It is doubtful whether the creation and publication of elaborate models for the preparation of the second type of prediction will be worthwhile in a world containing socially determined events, since it may be impossible to obtain adequate data for the testing of alternative models before the adaptive behaviour of economic actors has rendered the models obsolete[14].

Vol. 5
1977

The researchers propose to assign an optimal prediction model to each possible reporting method. They will then choose between the alternative optimal combinations which are available on the assumption that high correlation between predicted values and actual values is the predominant attribute of a desirable combination.[9]

It is not necessary that the relationship used in a prediction model should be explicable by present knowledge, but it is helpful if the relationship *is* explicable, since understanding will add credence to the prediction method and make it less likely, for example, that a spurious correlation would be accepted as valuable.

A crude linear extrapolation of a time series of operating cash flows would be unlikely to succeed as a practical prediction method because future results could depend significantly on various exogenous variables which could not be dealt with implicitly [15]. Operating cash flows could instead be estimated from turnover, by first creating independent predictions of turnover based upon

[9] An empirical test of feasibility would then precede practical implementation of any proposal for change.

general economic trends, and upon trends for the industry in which the firm operates. Many of the inputs for such a model may be supplied from sources other than the firm itself. The problems which arise in considering a claim that the highest predictive ability would be an attribute of the measurement method which gave the highest correlation between predicted future cash flows and actual cash flows may be summarised as follows:

(i) The problem of the limited lifetime of a rigid model for the prediction of future economic events and sequences of such events in a changing world, and

(ii) The problem of limiting the growth and complexity of a prediction model using several sources of information *within financial constraints* to something which must clearly be rather less than a full scale replica of an ideal corporate budgeting system.

THE SUBSERVIENCE OF PREDICTIVE ABILITY TO AN ECONOMIC CRITERION

Consider again the general claim (now paraphrased) that predictive ability is an overriding and justifiable criterion for use in assessing the optimality of competing accounting standards in meeting the decision-taking objectives of the users of accounts. This statement appears to result from an oversimplification of the multiple qualities which contribute to making the output from some accounting information systems more desirable than the output from others. If any single criterion must be used to distinguish between accounting methods which are equally acceptable on logical or *a priori* grounds, then 'usefulness' appears to be a stronger candidate than predictive ability. Clearly, the standard method which is adopted in any sphere of accounting should be that which makes the resultant accounting report most 'useful' to its users. Usefulness has been a vague concept for many years and examination shows it to be a complex one. One attempt [27, p. 231] to clarify the concept has identified some of the attributes of a measurement which contribute to making that measurement a useful one. They are:

(i) Relevance to the needs of its user.
(ii) Reliability (freedom from bias, and verifiability—which means, for example, that erroneous action should not be encouraged).
(iii) Understandability (consistency with user concepts, comparability and simplicity).
(iv) Significance (i.e. will affect the user's decision).
(v) Sufficiency (i.e. whether additional information from other sources will be needed for the user to take his decision).
(vi) Practicality (i.e. whether the measurement is timely and worth more than its cost).

Similar characteristics of 'useful' corporate reports were outlined in 'The Corporate Report' [28, p. 28–29]. The Inflation Accounting Committee has also considered the desirable characteristics of accounting information. Their report recognised that the attributes of objectivity, realism, prudence, comparability, consistency, intelligibility and ease and economy in preparation, are, to some extent, mutually inconsistent, and suggested that a satisfactory compromise between them has to be worked out [23, p. 62–66]. Some accounting systems are likely to involve more expense than others and this must be borne in mind in considering their comparative advantages. The net value of the accounting system to the larger system of which it is part involves calculation of both the cost of information and its value in improving system performance.

We recognise the extreme difficulty of measuring the incremental value of an accounting report to an economic system where predictions used within the system are based on information from multiple sources and not just from accounts. William Morris [22, p. 483] writes as follows:

> 'To apply statistical decision theory to information decisions, it is necessary to be able to enumerate the possible outcomes of future data collection efforts, and further, to compute the probabilities of these outcomes. In addition it must be possible to indicate just how the information will quantitatively change the decision maker's view of his choice.'

Vol. 5
1977

We are far from achieving Morris's targets for the application of statistical decision theory to the choice of accounting method since we have not yet described the information image of a user of accounts in practice or the method of its updating. Both of these things must be done before we can determine how accounts may alter the image and affect decisions. The analysis of the marginal contribution to value from an individual item of data within a report must be even more difficult than dealing with the effect of a whole accounting report, since we cannot know all the users who may find an individual item of data relevant or satisfying to one or more of their needs.

This paper accepts the argument that a pre-requisite for any set of accounts to be useful (i.e. valuable) to a decision taker is that it has some quality of predictive ability in relation to some aspect of a decision model he is using. We agree that the evaluation of reporting suggestions should proceed by incorporating the suggested information in financial models and examining whether the resulting decisions are superior to those reached without this information. We point out, however, that information has the potential to affect *all* the components of a financial decision model (i.e. actions, outcomes, probabilities and utilities) and not just to provide for the assessment and revision of the probabilities of outcomes which the decision taker looks upon as in some degree possible. An information system has a role in the construction of a model since the set of relationships comprising the model is indicated from experience. An information system has a role in the testing of the model because its output indicates

the adequacy of the model's assumptions and implications (predictions) and suggests required modifications. An information system is also important to the model throughout its life because it produces feedback control information to monitor the performance of the model and signal the requiring modifications in parameters and variables as the environmental conditions change [18, pp. 101–102]. The desired level of complexity of a model is a practical matter, and will depend upon the rate of change in environmental conditions and the incremental value from greater complexity. These multiple roles of information systems in affecting the components of financial decision models should not be neglected when the interrelationships of the accounting report (as a model of the entity) and the financial decision model are being considered.

Effective use of a criterion of value is not easy. For example, the value attached to accounting information for the investor in ordinary shares will depend upon the scope of the system which is surveyed.

The production of information generated purely for trading purposes within a large stock market may appear privately advantageous, but it is likely to be socially suboptimal, since it requires real resources and redistributes investor wealth without creating it [11]. From a social point of view, if we are to propose any changes, we must establish that different accounting information would aid the establishment of a set of security prices, such that there would exist an optimal allocation of resources among firms and an optimal allocation of securities among investors.[10] We must also be convinced that the incremental value to society from the new allocation exceeds the incremental information cost. The appropriate overriding criterion for change will *not* be that an alternative information system provides better inputs to a prediction model but that a system in which we are interested will be 'more valuable' in contemplation of change to the new or amended information system than without such change.

REFERENCES

1. Arnold J (1973) *Pricing and Output Decisions*. Accountancy Age Modern Finance Series, Prentice-Hall, Englewood Cliffs.
2. Arnold J and Hope A (1975) Reporting business performance. *Acc. Bus. Res.* **18**, 96–105.
3. Ball R (1972) *Changes in accounting techniques and stock prices. Emp. Res. Acc.* Selected Studies, Supplement to *J. Acc. Res.* **10**, 1–44.

[10] This may be difficult, not least because of the empirical evidence that large stock markets already make efficient interpretations of economic events from accounting reports despite the variety of conventions by which they are currently represented. For examples see [3 and 17]. Ijiri *et al.* [16] reach similar conclusions in relation to the ability of management to recognise economic reality from accounting measurements despite the different reporting modes. They conclude that a decision change will only result from an accounting change where there is no feedback on the performance of the accounting process (or during the time lag before the feedback loop is completed) and where the accounting change is ignored because of functional fixation.

4. BEAVER W (1966) Financial ratios as predictors of failure. *Emp. Res. Acc.* Selected Studies, supplement to *J. Acc. Res.* **4,** 71–111.
5. BEAVER W (1968) Alternative accounting measures as predictors of failure. *Acc. Rev.* **43,** 113–122.
6. BEAVER WH, KENNELLY JW and VOSS WM (1968) Predictive ability as a criterion for the evaluation of accounting data. *Acc. Rev.* **43,** 675–683.
7. CARSBERG BV, ARNOLD J and HOPE A (1975) Predictive value: a criterion for choice of accounting method, unpublished.
8. CARSBERG BV, HOPE A and SCAPENS RW (1974) The objectives of published accounting reports. *Acc. Bus. Res.* **15,** 162–173.
9. CHAMBERS RJ (1966) *Accounting, Evaluation and Economic Behaviour.* Prentice-Hall, Englewood Cliffs, NJ.
10. CHAMBERS RJ (1975) A critical examination of Australian accounting standards. *Abacus* **11,** 136–152.
11. FAMA EF and LAFFER A (1971) Information and capital markets. *J. Bus.* **44,** 289–298.
12. FRANK W (1969) A study of the predictive significance of two income measures. *J. Acc. Res.* **7,** 123–136.
13. GONEDES NJ (1972) Efficient capital markets and external reporting. *Acc. Rev.* **47,** 11–21.
14. GREENBALL MN (1971) The predictive-ability criterion: its relevance in evaluating accounting data. *Abacus,* **7,** 1–17.
15. HOPE A and SCAPENS RW (1975) Rising prices, financial statements and fundamental research. *Omega* **3,** 267–282.
16. IJIRI Y, JAEDICKE RK and KNIGHT KE (1966) The effects of accounting alternatives on management decisions in *Research in Accounting Measurement* (Eds. JAEDICKE RK, IJIRI Y and NIELSEN O). pp. 186–199. American Accounting Association.
17. KAPLAN RS and ROLL R (1972) Investor evaluation of accounting information: some empirical evidence. *J. Bus* **45,** 225–257.
18. LEV B (1974) *Financial Statement Analysis: A New Approach.* Prentice-Hall, New Jersey.
19. MACE JR (1974) *Management Information and the Computer.* Accountancy Age Modern Finance Series, Prentice-Hall, New Jersey.
20. MACE JR (1977) A systems approach to the analysis of financial reporting. *Acc. Bus. Res.* in press.
21. MAFFEI RB (1958) Simulation, sensitivity and management decision rules. *J. Bus.* **31.** 177–186.
22. MORRIS WT (1964) *The Analysis of Management Decisions,* Revised Ed. Irwin.
23. REPORT OF THE INFLATION ACCOUNTING COMMITTEE (1975) Cmnd 6225 HMSO.
24. REVSINE L (1971) Predictive ability, market prices and operating flows. *Acc. Rev.* **46,** 480–489.
25. REVSINE L (1973) *Replacement Cost Accounting.* Prentice-Hall, New York.
26. SIMMONS JK and GRAY J (1969) An investigation of the effect of different accounting frameworks on the prediction of net income. *Acc. Rev.* **44,** 757–776.
27. SNAVELY HJ (1967) Accounting information criteria. *Acc. Rev.* **42,** 223–232.
28. THE ACCOUNTING STANDARDS STEERING COMMITTEE (1975) *The Corporate Report.*
29. THEIL H and LEENDERS CT (1965) Tomorrow on the Amsterdam stock exchange. *J. Bus.* **38,** 277–284.

ADDRESS FOR CORRESPONDENCE: *JR Mace Esq, Dept of Accounting and Finance, University of Lancaster, Gillow House, Bailrigg, Lancaster LA1 4YX, UK.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

Investment and the Swedish Economy

HG JONES

University of Strathclyde, Glasgow

(*Received February* 1977; *in revised form March* 1977)

The causes of the high rate of investment in the Swedish economy are discussed with particular emphasis on the relevance of restructuring; however, restructuring may be more difficult in future. The source of funds and the function of the Investment Bank are described and the debate on the scale of future investment is summarised.

THE ECONOMY of Sweden has for several decades exhibited a steady growth and achieved, by 1974, a level of GDP per capita surpassed by only one country in the world. The industrial scene has been dominated by change and by a willingness to accept change: there has been a strong consensus of opinion that what is good for Sweden is good for the man in the street. In recent years, there has been an increasing desire to identify the labour force with the firm. In marked contrast, with this high growth, dynamic and largely consensus-based society, many other economies have been riven by internal dissension arising from conflicting interests and have been able to achieve only slow growth.

With this background, it is appropriate to examine the Swedish economy from many angles in the search for indicators for the less fortunate; this paper is concerned in this spirit with certain aspects of capital investment.

The gross investment in Swedish industry in 1975 reached an all time high of nearly £2400 per employee (at current rate of exchange).

This figure was achieved in spite of four factors which might in other countries have been expected to militate against investment. Firstly, there is widespread concern that the new labour laws [7] becoming effective in late 1976 or early 1977 will impair the profitability of many companies and even impose threats to the survival of some. Secondly, there is just as much concern that the rising wage and social costs of employment will impair the competitive capacity of Swedish industry, and hence its profitability: Table 1 illustrates the position for the textile industry. Thirdly, since 1960 there has been a remarkably steady trade cycle of 4 years in Sweden and the period 1975/76 would be a period of recession on this basis (and in many respects, e.g. output of steel [3], it has been). Fourthly, between 1960 and 1975 the output per man-hour in manufacturing industry

TABLE 1. ADDITIONAL EMPLOYMENT COSTS IN THE TEXTILE INDUSTRY PER MAN-HOUR IN SWEDEN RELATIVE TO OTHER COUNTRIES (in S.kr.)[8]

	1972	1973	1974	1975
USA	0·20	2·20	3·75	7·45
UK	8·80	10·80	11·45	15·70
Portugal	14·30	16·20	16·20	17·50

has increased at an average rate of 6·6% compound per annum, an enviable result but one which has required considerable restructuring of industry: while restructuring has not ceased, it will become more difficult in future partly because the changes already achieved leave less room for manoeuvre, and partly because social influences are likely to resist substantial restructuring in future years.

In the Swedish business philosophy, the first three factors instead of reducing investment are reasons for increasing investment because this seems the natural way of overcoming the difficulties they pose. The outcome (and of course not only for the reasons cited) has been that the number of people in employment has risen steadily by a total of nearly 6% over the years 1973–1976 while the unemployment figure was only about 1·7% throughout 1975 and 1976, the peak being 3·0% for one quarter in 1971. The general index of share prices on the Stockholm Stock Exchange has also risen fairly steadily by a total of 125% between 1971 and 1976. These results would indicate a successful economy, particularly when viewed against world performance, even though there has been little increase in GDP for 1975 and 1976.

Vol. 5
1977

RESTRUCTURING

The dramatic type of restructuring is that of a firm making a quick decision to move out of a comparatively unprofitable product line into a more profitable one, usually within the same industry. This practice is fairly widespread in Sweden and implies management of a high order both in assessing the position and then in taking effective action. Examples which might be quoted are the well publicised decision by Kockums to specialise in the building of modern tankers, the less well known action of Gummifabriks in pulling out of tyre manufacture, and the more painful (since it involved major cuts in employment) operation in Algot Johansson to rationalise product lines by concentrating in Sweden on casual and sports wear. Each of these very different actions has involved capital investment in plant or equipment to streamline production methods without which the restructuring would have been pointless.

Another type of restructuring has been brought about by the concentration

of production in larger units in order to obtain the economies of scale. The conspicuous examples have been in steel, pulp mills and chemical plant, in all of which some rationalisation of products has occurred at the same time. The policy has been successful with special steels in which Sweden still has a substantial export market of about £400m pa, the emphasis being on high value products produced by new processes directed towards maintaining high standard of purity and homogeneity, or, in the words of Uddeholm [10] "it is not only the steel itself that we sell today. We often also sell solutions to customers' problems". With special steels, quality is all important and the production units, although say three times the size of a decade ago, are still not so large as to be beyond the resources of a small country or of a single company.

The position is different in mild steel where plants producing in excess of 5 m tons pa are required for the benefits of scale to materialise. This size is beyond the capacity of the private companies (three main works each around 1 m tons pa) while the experience of the still small state owned industry in the North has been none too happy. It is cheaper to import mild steel and the debate now centres on whether Sweden can afford a mild steel industry; for strategic reasons—economic as well as military—production will probably continue. It looks as if the development programme in the North will be much less ambitious than the original "Steelworks 80" project. The argument is that it would be better to base developments on an existing site such as Oxelösund where the expertise and steel making traditions could be a helpful factor. Such a restructuring would require a consortium of the private companies and the state, and looks far from impossible.

Restructuring of the shipbuilding industry seemed likely in 1976. Two major companies approached the government for financial assistance in exchange for shares; Kockums is in fact now the only major shipyard in Sweden not in some form of state ownership and since it is profitable, the company wishes to remain independent. Two points are of interest, namely, the process of restructuring (followed by rationalisation) is voluntary, and secondly that Kockums objects not on ideological grounds but because of economic considerations—if any economic advantages to the company was apparent, Kockums would be prepared to review the position.

A third factor in the restructuring process is partly a consequence of the other two: it is the disappearance of many small companies. In 1973, manufacturing industry was made up of 2000 (or 14%) less establishments than there were seven years earlier although the total number employed therein had marginally increased. Mergers, takeovers or sometimes just financial failures had all taken their toll of the small companies. When one sees in Table 1 the way in which labour costs have been rising more rapidly than in other countries, the problems of the small company can be appreciated. Lacking a sound capital base, the small and often labour intensive companies have not been able to install sophisticated machinery as a partial replacement for labour.

The restructuring has laid heavy claims on the capital resources of the country. It has also involved heavy social costs because it has required occupational and geographical mobility on the part of labour. There have been great changes in the structure of the labour force; for instance, in a decade or so, the number of people working in textiles, farming and forestry has almost halved while engineering has increased its manpower by 10%. In the past, these costs have usually been met without protest because it has been the policy of the central body of the blue collar unions (LO) to support high productivity through modernisation as a means of obtaining high wages. The high standard of living demonstrates that in material terms the policy has undoubtedly paid. The Labour Market Board, too, has helped by being geared to assist in retraining and in providing migratory allowances, normally for workers moving from the rural areas to the towns.

The pattern could well change in the future. There is a strong 'green movement' that seeks to reverse the migratory trend by extolling the virtues of life in the country. The Labour Market Board is now more concerned to steer industry, by means of grants and training allowances, and by a system of permits, to the people wanting jobs in the old 'deprived areas'. LO has behaved as an authoritarian body exercising a strong central control over the rank and file; in order to obtain a satisfactory outcome from the new law on co-determination in industry, LO are now proposing decentralised decision making at local level. It is less easy at local level to take the broad view on what is good for the country if one's own job is at stake because of restructuring or for any other reason. In the past, LO has been strongly in favour of industrial rationalisation, and overmanning has been no part of its brief; there are now signs for the first time in Sweden for many years of job preservation becoming an issue.

Vol. 5
1977

These are major changes in outlook and cumulatively will certainly make the continued modernisation of Swedish industry more difficult. For other reasons, the rate of structural change has already slowed down, the rate in the first 5 years of the 1970's being only half of that in the 1960's, during which some economists believe two thirds of the increase in productivity was due to restructuring. The question then is, can Swedish industry continue to be profitable? Without profits, continued modernisation through investment is difficult and the competitive edge could be lost.

A further factor not helping profitability is the move from incentive payment schemes to monthly wages with little or no incentive element. Pressure for the change was strongest in the North where in forestry and iron ore mining, climatic conditions could make it difficult and unpleasant to maintain bonus earnings at a constant level. In these areas, with fixed wages losses in productivity at times up to 20% have occurred due to late starting, early finishing, stops for chats and less urgency in rectifying halts due to machine stoppages; a suggestion of overmanning is also creeping into the situation. Elsewhere only

some companies have changed to monthly payments: a few recognising the problem have endeavoured to combat this industrial malaise by creating job satisfaction and interest in the Company.

TABLE 2. PERCENTAGE RETURN ON ACCUMULATED CAPITAL IN PLANT AND EQUIPMENT FOR SELECTED INDUSTRIES IN SWEDEN [2]

	1974	1965	Difference 1974–75
Chemical	22	15	+ 7
Mining	17	31	−14
Cars, lorries, aeroplanes	16	11	+ 5
Rubber products	6	8	− 2
Textiles	6	10	− 4
Shipbuilding	5	0	+ 5
All industry	13	10	+ 3

(Before tax but after depreciation)

In case the foregoing seems unduly pessimistic, Table 2 is quoted to restore the perspective; the last line indicates that in 1974, Swedish industry showed an average rate of return on total capital employed of 13%, a better figure than in 1965 and better than the 10·3% average of the years 1970 to 1974. (An after tax return on "total capital for total industry" of 5·1% is quoted for 1974) [1]. Of the arbitrary selection of industries, textiles and rubber products are both in a declining position in spite of strenuous efforts to halt the trend. Chemicals and transport show a good and improving situation. The special problems in the mining industry, mostly for iron ore, were a sharp drop of about 50% in the world price for iron ore relative to other commodities, due to competition from new ore fields, coupled with rising costs due to the need to go deeper into the mountain and also due to the demand for better working conditions; in spite of a return of 17% there is little room for complacency, because the benefits expected from massive injection of capital have still to be realised.

A rough measure of the state of industry and of the degree of investment is the proportion of equipment less than 10 years old: for the combined mining and manufacturing industries in Sweden, this has risen from 61·4% in 1965 to 65·8% in 1975. [4] The figure is even higher for forestry—86%—a reflection of the changing practice of recent years.

SOURCE OF FUNDS

Swedish industry continues to self finance investment through the Investment Funds which enable 40% of the pre-tax profit to be set aside tax free for future

investment in plant, equipment and stock building (46% of the money must be deposited in a non-interest bearing account in the Central Bank). The investment has to be approved by the Labour Market Board which thus exercises some control over the economy, both in the location and timing of expansion. This system has the merit of keeping capital in the profitable sector of the economy, as does also the arrangement under which a company has the right to borrow back 50% of its contribution to the National Pension Funds.

Self financing accounted for 75% of industrial investment in 1974 but only 50% in 1975. Thus to finance modernisation and expansion on the level necessary to sustain the expected growth in GDP, in addition industry has to resort to bonds and bank loans, for which the interest rates are still less than 10%. A new company in particular does not have profit and may not have contributed to the Pension Funds so the first approach will probably be to a bank. The first and most important criterion in at least two banks when deciding whether the loan will be granted is the quality of the management: the state owned Investment Bank even goes so far as to visit the company to investigate the management structure before granting a major loan. The second criterion is the track record of the applying company. It is not until the third stage that the project itself becomes an issue. The Investment Bank also insists on adequate market research before granting a loan, and if necessary will either itself undertake this work or hire specialists so to do. A return on capital of about 20% is sought by all the Banks but since one function of the Investment Bank is to help finance new and possibly risky ventures of promise or alternatively, very large projects, a return lower than this may be acceptable if the project seems in the country's interest.

Vol. 5
1977

In practice, the Investment Bank in most cases tries to act as a 'topping up' agent to ensure a project does not suffer from under financing; that is, a company is expected to secure most of its capital from other sources before the Investment Bank adds its quota. A new venture may not be able to get other support, in which case the Investment Bank may act alone. If its capital is at risk, the Investment Bank has no hesitation in stepping in as 'company doctor' either to cure the ills through competent management or reluctantly to kill the patient as painlessly as possible. Although profit making was not the prime reason for the establishment of this bank, it has in fact secured a return on total assets of 1·5% in 1974 and 1% in 1975 [6]: the falling return is a reflection of the problems of small companies and of the difficulties in 1975.

From the nature of the enterprise, not all of the new ventures can be expected to yield financial success but in fact losses in individual projects in 1975 amounted to only £4000: some have criticised the Bank on the grounds that it is not losing enough! The Bank profits indicate that a high success rate *is* achieved due largely to the capabilities of the personnel in 'sniffing out the winners'. This process is probably helped by the existence of two government sponsored bodies, namely, STU with about £45 m available pa to support development up to prototype

stage and which therefore has some research involvement, and STUFO with about £4½ m pa to support the transition from prototype to small scale production in order to test the market. With both agencies, the grant has to be repaid if the project is successful. Arne Callans, the Managing Director of the Investment Bank, is Chairman of the latter agency and is thus in a strong position to assess new projects at an early stage. This total combination is typical of the Swedish will to succeed.

Recent years have seen the emergence in Sweden of 24 associations made up of small businesses in the same geographical area combining together in order to secure specialist services in such matters as taxation or marketing. The Socialist government, which in the past had been guilty more of neglect of than antipathy towards the small companies, at last recognised their importance to the economy by granting loans to the associations of up to £40,000 in 1975. The Investment Bank, too, is now involved in loans of up to £60,000 pa to these associations, a development which is regarded as important, since it offers credit facilities to the small companies to a greater extent than ever before; in 1975, the Bank involvement was in excess of £2 m. A new feature is that the administration of these loans resides with the local association.

The Swedish Investment Bank takes a very positive role in encouraging new commercial ventures, preferably those with an innovatory element. Established by the government in 1967 with £100 m credit but no other resources, it has steadily built a reputation for success so that by 1976 it had about £600 m in outstanding loans of which about £235 m were new commitments in 1975, spread over 142 companies in a wide range of industries. And all of this together with a major asset of a magnificent new office and shopping complex in the heart of Stockholm achieved with a staff of less than 40 (mostly economists but with a sprinkling of other disciplines). An indication of the Swedish way of government is that throughout the 9 years of the Bank's existence there has been no interference by the cabinet or by ministers.

As a result of taking over ailing companies, the Bank now has 90% or more ownership of four sizeable manufacturing companies with products ranging from die castings to readymade clothing. The tone of activities can be appreciated from remarks in the annual report that "the restructuring of the new company has taken approx 15 months during which a new product was developed, a production apparatus built up and a marketing function organised". In case this quotation leaves the impression of ruthless hatchetmen at work, it should be said the atmosphere in the Bank is one of relaxed competence. With a success rate of about 1 in 5 in the rescue area, not all of these companies are expected to enter a profitable phase, but that is not for lack of effort. The recovery process is assisted by a wholly owned consulting company (with 8 specialists), 85% of whose activities are in support of the Bank's interests.

FUTURE INVESTMENT

Various projections of the rate of growth of GDP in Sweden have been made, all supported by substantial documentation built up industry by industry. The official survey of the Ministry of Finance [5] quotes a maximum growth rate of 6·2% pa for the production of goods by industry for the period 1975 to 1980; this is one of four alternatives under examination, the lowest having a growth of 4·7%. These alternatives are all less than the very considerable growth rate of 7·5% achieved in 1960 to 1965, but in view of increasing difficulties appear optimistic when viewed against the growth of 5·1% or less expected for 1970–1975.

The Industrial Institute for Economic and Social Research (IUI) believes figures of 4·1% or 5·4% pa for industrial production (2·9 or 3·6% [11] for the growth in GNP) are more realistic, the alternatives depending on the distribution of funds between the public and private sectors.

On important issues, Sweden has for two decades at least taken the view that a second independent investigation to supplement or to balance the official report is very desirable, and this accounts for the prominence given to the IUI Report. Before arriving at the growth figures, this survey analyses the influence of devaluation as a possible strategy which it describes as "a most dubious method" of improving the competitive position because it can only lead to a spiral of rising wage and price costs, and in so doing discourages investment: the UK can demonstrate the validity of this hypothesis! The report also gives serious consideration to revaluing the Swedish Krona in line with the German Mark, but concludes that Sweden must first demonstrate an ability to limit rising production costs. These are very frequent talking points in Swedish management circles at present with vocal support for both policies.

The IUI report emphasises that increase in real wages can only be obtained by means of increased productivity in private industry, which must support wage increases in the public sector also: coupled with this is the need for intangible investment in the development of new products and the creation of new outlets in overseas markets. Without investment increasing at about 7% per annum (for the more optimistic growth in GNP), rise in productivity is impossible and without 'considerable increase' in profits, investment is unlikely. If investment cannot be obtained any other way, then financing through taxation is not out of the question so that wage earners would be repaying part of their wages to the companies via taxes—a novel but true way of regarding investment grants. All in all, a domestic inflation rate of 7·5% pa is predicted for Sweden compared with an international rate of 6% in the IUI survey (or 3% assumed in the official report). Lack of wage restraint could place great strain on the economy and unemployment would result. It will be interesting to watch the performance of the new Moderate Swedish government in the pursuit of these economic policies which presumably it will favour rather than the more

Vol. 5
1977

expansionist assumptions of the official document inspired by the outgoing Socialist government.

In order to attract sufficient investment to achieve the growth rates of the official programme, the Waldenström Committee set up by the Employers' Confederation (SAF) in 1974, claims [9] that profit levels will need to be at least doubled.

The Committee also recommend private savings should be encouraged because these can be a useful source of funds. The Committee is strongly against the Meidner type of funds proposed by the unions, in which 20% of all profits would be transferred to employee controlled funds—demands for these funds were advanced for the double purpose of influencing decisions affecting the worker and of transferring economic power to the unions. The Waldenström Committee regarded this proposal as being equivalent to the compulsory confiscation of assets owned by the private shareholders, and would make it more difficult to attract new capital to industry. Also instead of being a minor reform to equalise wealth as the unions claimed, the employers maintained the Meidner Funds would lead to a drastic transformation of the economic system. But the committee is not opposed to voluntary employee funds based on about 1% of wages and with the contributions frozen in the funds for say the first 7 years; these funds would purchase shares in the usual way, and if they attract tax concessions, then the funds should be open to subscription by private citizens as well as to particular wage earners.

With the change in government, action on the Meidner Funds appears to be in abeyance for the time being. There is little doubt that substantial investment from some source will be required if the growth in Swedish prosperity is to be maintained.

ACKNOWLEDGEMENTS

Thanks are due to many Swedes for discussion on the subject matter of this paper, and to the Swedish Institute for financial assistance.

REFERENCES

1. Åkerman A (1976) Can small firms survive? *Skand. Enskilda Banken Q. Rev.* (2), 40–46.
2. Anon (1976) "I.U.I's långtidsbedömning 1976. Utrecklingsvägar för svensk ekonomi fram till 1980." I.U.I. Stockholm.
3. Anon (1976) Inrikes Ekonomi. *Skand. Enskilda Banken Rev.* (6), 22.
4. Anon (1976) *The Swedish Economy* 1975–1980, p. 114. Ministry of Finance, Stockholm.
5. Anon (1976) *The Swedish Economy* 1975–1980, p. 189. Ministry of Finance, Stockholm.
6. Callans A (1975) Annual Report. Sveriges Investerings Bank A.B. Stockholm.
7. Jones HG (1976) Sweden's changing tides. *Mgmt Today* June, p. 80.
8. Söderholm O (1975) Annual Report. Algot Johansson A.B. Borås.

9. WALDENSTRÖM E (1976) Report on Company Profits, sources of Investment Capital and Wage-Earner Funds, pp. 1–10. SAF Stockholm.
10. WESSMAN G (1975) Annual Report. Uddeholm A.B., Hagfors.
11. WOHLIN L (1976) Competitiveness, the Balance of Payments and the Allocation of Resources: A Long Term Appraisal. I.U.I. Stockholm.

ADDRESS FOR CORRESPODENCE: *HG Jones Esq, Senior Research Fellow, Dept of Operational Research, University of Strathclyde, Livingstone Tower, 26 Richmond Street, Glasgow G1 1XH.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

The Value of Information Concept Applied to Data Systems[1]

DONALD V MATHUSZ

Directorate of Combat Operations Analysis, US Army Combined Arms Combat Developments Activity

(*In revised form March* 1977)

Cost–benefit analysis has a considerable literature in which information systems have been patently ignored. This reflects the considerable difficulties of applying the theory to information systems, and the state-of-the art remains relatively as Koopmans described it some 19 years ago (1957). A bar to further development would appear to be the lack of an applicable value-of-information concept. This paper seeks to clarify the issues and provide a robust theoretical and data analysis framework that will cover most situations. The approach here is to separate explicitly the dimensions of cost from those of information benefit, and examine the implications. The Null Information Benefit condition emerges as a special theoretical case, but potentially a most important one in applications. This case together with the Pareto optimum defines a large class of such problems that can be handled by the decision criteria and data analysis techniques tabulated and discussed here. The selection of input data techniques defines the limits of later project justification and may be crucial to the political viability of the projects throughout its life. Finally, the general management vs information systems management relationships are discussed in terms of this situation.

1. INTRODUCTION TO THE STATE-OF-THE-ART

"If the cost of information processing is to be balanced against its contribution to decisions, a more general method for quantifying the several relevant aspects of the intuitive notion of information is needed."

T. C. Koopmans—1957

THE THEORY of cost–benefit analysis is essentially incomplete in many respects, e.g. where it may involve a social welfare function [33]. The more common cost–benefit analysis of hardware systems is also by no means without its pitfalls [13, 15, 21, 22]. It seems only natural then, that computers have been turned to for what aid and comfort they can yield but with very limited success

[1] The contents of this paper represent the views of the author and should not be considered as having official Department of the Army approval, either expressed or implied.

[31]. It can be argued that there is a considerable evaluation potential in the man–computer interaction schemes as demonstrated in MIT's project MAC that could be applied to the cost–benefit problem, perhaps via on-line simulation [19, 27]. However, it is not the lack of computing power, but lack of a framework in which to put the computing power that is the essence of the problem.

In these analyses, the Management Information System (MIS) itself has been seriously neglected. Part of the reason may be the considerable confusion that exists as to how cost–benefit analysis is applicable to the MIS. Opinion ranges from that of advocating profit and loss (P&L) evaluations, to those that would not attempt any cost benefit analysis. At least, this latter view is implied by detailed expositions of the potential of MIS without mention of quantitative benefits or costs [10, 11, 12]. McDonough advocates setting up successive implementation bench marks versus expected costs, by which cumulative costs can be plotted against progress for management control purposes [30]. This still leaves unanswered the question of whether the project should have been undertaken at all.

Historically, justification of computers and their software as investments has followed two paths. In the first, the computer has been viewed as an overhead or general service to the organization. Its justification then was on the general grounds of added capability. This seems to have come about from the early wartime uses made of the computer by the scientific community. Under these circumstances, the only economic analysis required on a computer support proposal was that of examining the magnitude of the expense. Management then decided whether it should be funded or not (sometimes government contracts helped).

The Profit and Loss (P&L) approach evolved from the computer's substitution for some tangible clerical operation. The first electronic accounting machines were early examples of this evolution which has proceeded to the present large computer systems used for payroll, billing, etc. These computer systems essentially produce the same tangible product as previous clerical systems with, it is hoped, a lower cost. In this instance, the investment justification is some variant of P&L analysis that will estimate the expected return on investment, such as profitability accounting [17, 18].

The MIS cost–benefit problem then reduces to: do we avoid the problem by assuming it away, i.e. by simply calling the MIS an intangible overhead service? Then, if we do this, what do we have left in the way of a comparative selection basis between candidate MIS projects?

This paper seeks to establish a workable cost–benefit approach to the selection of candidate MIS projects, in which the methods and techniques are those that corporate management would ordinarily expect to see in project justifications. Aspects of the subject which might be theoretically useful but esoteric to most management, such as Bergsonian utility, have been omitted [2].

2. PRELIMINARY CONCEPTS

The cost–benefit analysis of MIS is by definition concerned with two streams of values,[2] i.e. the information stream and the costs stream. Because there are conceptual difficulties with the former, the entire analysis oftentimes falls by default into the costing area [23]. Here, frequently we can devise a computer system with net negative costs, i.e. cost avoidances which are greater than the implementation and operating costs together. This money savings is then considered the benefit in practice. While this is not entirely in error, in the sense that it is a portion of the benefit, it neglects the information–benefit side and implies certain assumptions about the information characteristics involved. How do we evaluate the situation, for instance, where we replace a manual once-a-year, sales forecast, with twelve computerized forecasts, one issued every successive month, for the next 12 months ahead? It should be obvious that the mechanized system represents a considerable increase in potential corporate capability. If the P&L analysis shows a loss, do we conclude that the project should be rejected on that basis?

2.1 *Partition of information and cost-benefits*

To carry the argument to its conclusion, if the new system will not improve the information in some way, e.g. the same reports will be generated with the same timeliness, then the marginal difference in information-benefit between the two systems is zero. We may then justifiably evaluate the system solely on the basis of comparative costs. It follows then, that whenever we use costs without consideration of information benefits, we implicitly assume the information-benefit to be the same for all the systems and the information-benefit differences between systems to be zero (null).

2.2 *The information-benefit concept*

The information-benefit notion is: as a business (military) manager, what is the benefit of having particular information? To crystallize the concept, let us assume that there is $1000 buried in a box somewhere. How much benefit does knowing the location amount to? Clearly, this depends upon the recovery cost. If it is somewhere in our back yard, so that the recovery cost is negligible the information value is $1000. In the same manner, any recovery that costs $800 reduces the information's value to $200, and so on.

The difficulty in applying this benefit notion to real MIS seems to be a compounding of two problems.

a. Each manager (decision maker) has an individual input information profile that corresponds to his observational and thinking habits [1].

[2] We use the word benefit in a very general sense and reserve the word value to imply that the benefit can be evaluated in a common measure, such as dollars in business or reaction time in the military.

b. Even if we accept subjective probability, there remains the problem which Vickery defines as uncertainty as contrasted to risk. In the case of risk, we can assign probabilities to known outcomes, but in uncertainty not all the outcomes are known [40].

As previously stated, if we restrict the project selection criteria to cash flow, we have assumed the information benefit problem away, by assuming the candidate MIS projects all yield the same value of information. This in effect, is a Null Information Benefit assumption. It should also be noted that the Null assumption does not necessarily require identical timeliness of information between systems.

It may be that the information is produced periodically, such that its usefulness is unaffected by the precise moment of its production within the period, provided only that it is available when conventionally needed, e.g. end of the month. The same argument holds for the information content, although this may be manager-dependent as noted by Ackoff [1]. Thus, the Null Information Benefit assumption may be valid under a far larger number of business applications than would seem likely in the first instance.

As an example let us consider the arbitrage situation. Here the objective is to buy and sell a given commodity in different markets, e.g. gold in Zurich and Hong Kong and profit by any price differences. The firm engaged in arbitrage operations does not necessarily wish to use the commodity itself, but profits by producing a price orderliness between markets (ideally). The value of the information is the potential profits implied in these price differences between markets. In this situation, the net potential profitableness is affected by the cost of producing the information (which should be fairly constant), and the expected commodity volumes and prices involved (fairly volatile). Here, the information's value can be estimated in terms of expected dollars. The timeliness of the information affects the commodity volumes that can be arbitraged during the active trading day. When the markets close however, the information on the final trading has constant value until the market opens again. Thus, in this simplified example, we would not pay more for a system that produced the final trading results seconds after the markets closed, as compared to one that yielded the same information an hour later, since both systems would allow us ample time to plan the next day's opening trades.

When the Null assumption clearly does not hold, the simplifying power of the Pareto optimum can be applied [2]. If one system has both lower costs and superior information benefits, it becomes the obvious choice. In the same manner, if both systems have equal costs, the one with superior information benefits is the obvious choice. This implies that systems with the same information benefit can be validly evaluated on costs alone. In some instances where neither costs nor benefits fit the Pareto criteria, it still may be possible to convert the information benefit into dollars. This is usually difficult for a MIS, but where it is possible, the evaluation is essentially then a costing process. In

Vol. 5
1977

those instances where neither the Pareto criteria is satisfied, nor is converting the information benefit into dollars possible, the final decision must be based on management experience as in the past.

2.4 *Simulations—a robust approach to secondary effects*

In the large firm there may well be expenses or savings in departments other than the MIS project proponent. These secondary effects may diffuse throughout the entire organization in a complex way. A model or simulation may be needed that will represent these interactions and summarize the effects in terms of aggregate cash flows. In those circumstances, simulations by digital computer have been useful, at least since the time of Forrester's basic book [14]. At this point, we would direct the reader's attention to Bonini's work, "Simulation of Information and Decision Systems in the Firm" [5]. Bonini demonstrates that simulations of this type can be done with results that appear to be very reasonable by the criteria most businessmen would use.

However, it is only reasonable to add that simulations require a high level of professional competence as well as a good deal of work. Under these circumstances, preanalysis is desirable to verify the need for a simulation, before any commitment is made to this approach [16, 32].

3. THE COST ANALYSIS

The first cost analysis of a project should serve only to establish the order of the costs involved [41]. We would include in this a feasibility investigation concerned with the most difficult implementation problems of the project and their impact on expected costs [25]. This aspect of the study might involve, hardware capabilities and limitations, software availability, operations research to define the system and data analysis to establish the input data requirements and their availability [26, 34, 37, 38, 39].

3.1 *The going concern analysis as comparative statics*

The going concern analysis is, by analogy, a kind of double snapshot comparison of the corporate economic situation. The question asked is: if the new system were fully installed and operating, what would the comparative cash flows be for some prior year, both with and without the system. Conventionally, this comparison period is the most recent prior year for which data is available. In essence, we are contrasting two economic situations, each of which is static in the sense that the corporate (agency) organization is in equilibrium once with the fully implemented system and once without it. The results of such an analysis might look like one period (year) of Fig. 1.

A natural extension of this approach, is to include a number of future periods (period analysis) in the comparison, but based upon projections of future work

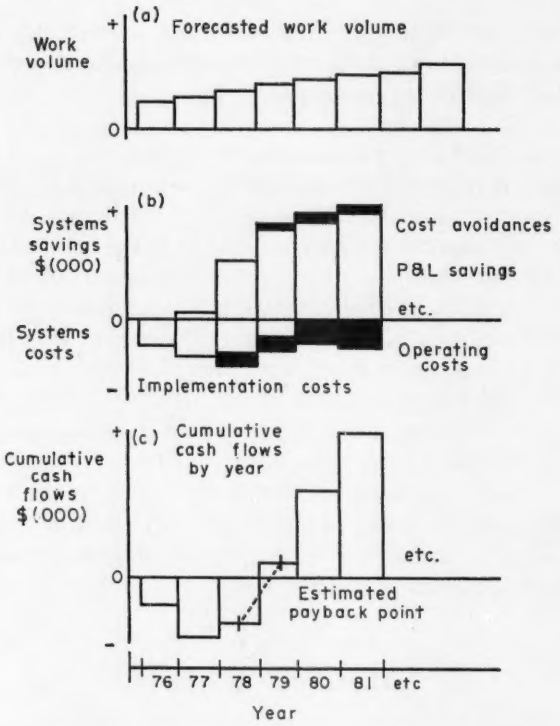


FIG. 1. *Estimated work volume, costs and savings and cash flows.*

volumes. Under these circumstances, total costs or savings may be separated into 'hard' savings based on current volumes and cost avoidances based on the projections of future increased volumes. Results of such an analysis might appear as Fig. 1b. It is worth noting that sampling techniques may be used to advantage to make the management information system relatively independent of increasing volume of input data for some applications [8, 24].

3.2 *System capitalization and cash flow dynamics*

Our static comparison gave us the before and after views without having to consider the implementation cost. If the results of this going concern analysis are satisfactory, it still remains to be seen if the development and implementation costs are justified.

The cash flow analysis is essentially a period analysis showing the combined results of the static and transient costs by period over time. Figure 1a, b, and c, show the development from data volumes into period costs, savings and finally into the cumulative cash flow. In large expensive systems, the cash flow by

Vol. 5
1977

period is critical in itself. These cash flows are needed by both corporate planning and the controller for investment feasibility analysis, while the chief executive will be chiefly interested in the payback and profit rates [4].

The question of the proper rate of interest is always raised. The answer most given is: if the funds are to be borrowed, then the loan rate is used; if internally financed, then the opportunity rate is used [29]. In practice, we suggest three or four rates be evaluated that include both the above rates. Then regardless of who challenges the rate *ex post*, we have figures or at least approximations to them immediately available. We can also use these results as a test of the sensitivity of the evaluation to the rate of interest, a feature often of interest to management.

4. MANAGEMENT VERIFICATION OF CASH FLOW ESTIMATES

The management of the project's cash flow should be based upon cash flow projections for which verifiable data can be obtained, i.e. data which can verify that the cash flow projections actually occur after the MIS project is implemented. This requires that certain actions be taken before and during the project. The advantages of this verifiable approach are twofold.

a. The system's manager is protected from arbitrary ex-post judgments of what the system should have and actually did do.
b. Top management may expect a distinct reasonableness of cash flow benefit claims.

This implies that the cost measurements will be taken with, or before the systems analysis is complete, and definitely before any system implementation. Some exception to this rule is possible, but only at an ultimately higher verification cost.

4.1 *Data techniques for the cost analysis*

Whenever we wish to change a system and measure the effects of the change on that system, we measure the system before (pre-state) and after the change (post-state). This concept underlines the basis of the theory involved in designated statistical experiments [9]. The techniques that are available to gather this measurement data may, for convenience, be divided into three classes, direct measurements, reconstructed data and models that synthetically represent the system over time.

The most straightforward of these techniques are those that measure directly each of the static pre- and post-states. Because these direct measurements are the most valid and lowest in cost, it is important to recognize that if this strategy is to be taken, the pre-state measurements must be made before any changes are implemented. As soon as the implementation of changes start, the system will

begin reacting and either be destroyed or badly distorted for measurement purposes.

There is, however, another group of reasonably simple techniques which would allow us to synthesize the system as it existed before the change and make estimates of the pre-state system. These generally cost more than the direct measurements. Also, this group will be found less useful for attempting to measure the system dynamics, since each measurement point on the dynamic path will require a reconstruction of the system at that point. If we really must synthesize this path, simulation would be a more appropriate approach.

Finally, there are techniques of obtaining data that we may need *ex ante*, e.g. sales forecasts. These are not techniques in the comparative statics sense, but rather what Hicks calls "Pure Positive Theory" [20].

4.2 *Requirements of a valid post-state*

In order to make any valid post-state measurements, we must have a stable system after implementation. This may be more difficult in practice than would first appear. To have this system, a minimum of the following conditions must be met.

a. Programs fully debugged and procedures firm: we propose that a systems program should not be considered debugged until it has operated successfully with real data for a minimum of three complete processing cycles. (This is an arbitrary but useful rule of thumb.)
b. Files fully cleaned, corrected and controlled: quality control processing may be built into the programs as exception routines, and/or quality control exercised on new inputs.
c. Personnel are experienced with the new debugged system over several full cycles of work—which may be in the order of a year or more, e.g. corporate financial reviews are usually done on a quarterly basis.
d. Stable organization—in order for procedures to be firm, the organization itself must be stable with a minimum of personnel turnover and no reorganizations.

4.3 *Reconstructed data, techniques*

a. *MTM Analysis* : MTM, or similar analysis in details of the old vs new systems to yield contrast man hours per unit workload [27].
b. *Standard data* : Utilizing standard data on previous workload via
   Personnel Standard Data, MTM Work Sampling or Position
   Machines Spaces [27].
c. *Vendor catalogues* : A search of old vs new catalogue/contracts prices to
   contracts establish contrasting costs.
   Total jobs
   Machines
   Materials

Vol. 5
1977

d. *Expert opinion* : While of doubtful validity, this approach may have the advantages of being the quickest, and lowest in cost—thus best for some rough or unique political situations.
Anything

4.4 *Direct measurement: techniques requiring a two state analysis*

Technique approach	*Pre-state analysis*	*Post-state analysis*
a. Vendor contract costs Total jobs Machines Materials	Establish requirements in detail such that contract bids can be sought any time.	Obtain bids on the old and new requirements to contract the total and per unit work costs.
b. Earned standard hours Personnel	Record workload, earned hours [7].	Measure unit workload and determine new standard and total and per unit costs.
c. Actual hours (recorded) Personnel Machines	Record workload and actual hours [7].	Measure new actual hours and workload. Contrast total and per unit costs.
d. Actual hours (unrecorded) Personnel Machines	Work sample over cycle and compute mean actual hours per unit workload and workload volume.	Work sample or record actual hours per unit and work volume. Contrast total and per unit costs.
e. Overtime: (situations of uneven high priority work) Personnel Computer rent Machine rents (Others)	Record overtime per workload volume and associated costs.	Record overtime per workload volume and associated costs. Contrast threshold work volume that produces overtime and unit and total costs.
f. Position levels: (Situations of significant overhead cost changes) Personnel	Record position structure as related to system workload and associated costs.	Examine position structure as required by the new system. Contrast total and per unit costs.
g. Position spaces: (full time activity) (situations of so-called non-measurable workload) Personnel	Record number of people doing this work and workload. Record associated costs.	Examine number of people required by new system. Contrast total and per unit costs.

5. MANAGEMENT INVOLVEMENT IN THE MANAGEMENT INFORMATION SYSTEM AND INFORMATION-BENEFIT

It is one of the standard laments of systems management that senior management does not become sufficiently involved in the creation of the management information system that is being designed to serve them. The problem that must eventually arise under these circumstances involves the credibility both of the information-benefit and cost estimates. This can become critical to the system's survival under the impact of the unusual internal political manoeuvering provoked by the interdepartmental nature of such systems.

While we see no easy solutions, the situation can only be helped by any definitive clarification to top management of the problems involved and by stimulating their involvement in the system's development. Stedry has made the candid observation that a cost control variance report is only as meaningful as the management's insight and action that results from it [36]. We have already noted that input information profiles change with managements. This has led to the concept of growing an information system about management [3]. The relevant point of this is that the capability to act on any given information is inherent in the manager himself. He must analyze his own management style and the input information profile this requires. He must then become involved enough in the system to ensure that his needs are met. It is nearly impossible for a given analyst to know *a priori* exactly what the system should give a particular manager. The manager's involvement is not only necessary but mandatory if the management information system is to be at all meaningful.

Vol. 5
1977

5.1 *Management risk and decision* (*with uncertainty*)

Management decision takes over where analytics stop (as was always the case in a well run organization). The technocrats, those in the Operations Research and Systems areas, must be candid enough to admit where the analysis ends, for any practical expenditures of time and money, and where management decision begins. To develop an honest rapport with higher management this aspect should be explicitly analyzed and presented for each specific project.

The technocrats must understand that even should their analytics be correct, there is still uncertainty to be taken into account (as defined before). One way to view the uncertainty is to consider that no analysis explicitly includes all causal factors, only the most significant of these. But it is possible that the non-significant factors omitted now, may contain one that will not only be significant, but even dominant in the future. For example, new legal interpretations of old laws may place a restraint on the organization in the future invalidating the entire post-state estimate. Many other surprises are possible and it is here that analysis must always be, in a sense, incomplete and management ability becomes

of paramount importance. Obviously, good analysis should decrease the surprise content of the future and become an adjunct to a well educated manager's decision criteria.

ACKNOWLEDGEMENT

The author is grateful for the encouragement and helpful suggestions of Dean K. Karger of Rensselaer Polytechnic Institute, who reviewed earlier drafts of this paper.

REFERENCES

1. ACKOFF RL (1962) *Scientific Methods: Optimizing Applied Research Decisions.* Wiley, New York.
2. BAUMEL WJ (1965) *Economic Theory and Operations Analysis.* Prentice-Hall, New Jersey.
3. BENETT E (1964) *Military Information Systems.* Praeger, New Jersey.
4. BEYER R (1963) Meaningful Costs for Management Action. In *New Decision-Making Tolls for Managers* (Eds BURSK EC and CHAPMAN JF). Harvard University Press, Massachusetts.
5. BONINI CP (1962) *Simulation of Information and Decision Systems in the Firm.* Prentice-Hall, New Jersey.
6. CAGAN C (1973) *Data Management Systems.* Melville, Los Angeles, CA.
7. CARROLL P (1954) *Time-Study for Cost Control.* McGraw-Hill, New York.
8. CYERT R and DAVIDSON H (1962) *Statistical Sampling for Accounting Information.* Prentice-Hall, New Jersey.
9. DAVIS OL (1960) *The Design and Analysis of Industrial Experiments.* Hafner, London.
10. DEARDEN J (1965) How to organize information systems. *Harv. Bus. Rev.* **43**(2), 67–73.
11. DESMONDE WH (1954) *Real-Time Data Processing Systems.* Prentice-Hall, New Jersey.
12. DUDLEY CL, JR (1965) Management audits of EDP installations. *Banking* **58**(3), 115–118.
13. EDWARDS NP (1964) On the Evaluation of the Cost-Effectiveness of Command and Control Systems. In *AFIPS Conference Proceedings.* Cleaver-Hume, London.
14. FORRESTER JW (1962) *Industrial Dynamics.* Wiley, New York.
15. FOX PD and HANEY DG (1961) Some topics relating to military cost-effectiveness analysis. Paper presented at the 29th National meeting of the Operations Research Society of America, Santa Monica, California.
16. GEISLER MA and STEGAR WA (1962) Determining preferred manager techniques in new systems through game-simulation. RM-3066-PR, RAND Corporation, Santa Monica, California.
17. GRAHM P (1965) Profit probability analysis of research and development expenditures. *J. Ind. Engng* **16**(3), 186–191.
18. GRANT EL (1950) *Principles of Engineering Economy.* Roland, New York.
19. GREENBERGER M (1965) The priority problem. MAC-TR-22, Sloan School of Management, Massachusetts Institute of Technology.
20. HICKS JR (1965) *Capital and Growth.* Oxford University Press, London.
21. KAZANOWSKI AD (1966) Cost effectiveness fallacies and misconceptions revisited. Paper presented at the 29th National Meeting of the Operations Research Society of America, Santa Monica, California.
22. KNORR K (1966) On the cost-effectiveness approach to military R&D. P-3390, RAND Corporation, Santa Monica, California.
23. KOOPMANS TC (1957) The Construction of Economic Knowledge. In *Three Essays on the State of Economic Science.* McGraw-Hill, New York.
24. MATHUSZ DV (1968) Data reduction through sampling in computer systems. *Software Age*, **2**(4), 21–23.
25. MATHESZ DV (1969) Software modularization. *Software Age* **3**(19–23); 39–43.

26. MATHUSZ DV (1972) Simulation input data analysis: some considerations and definitions. *ACM SIGSIM Q* 3(3)
27. MAYNARD STEGEMARTON and SCHWAB (1948) *Methods-Time Measurement*. McGraw-Hill, New York.
28. MEADOW CT (1973) *The Analysis of Information Systems*. Melville, Los Angeles, California.
29. MEADOW CT (1970) *Man–Machine Communications*. Wiley, New York.
30. MCDONOUGH AM (1963) *Information Economics and Management Systems*. McGraw-Hill, New York.
31. NELSON AC *et al.* (1967) Evaluation of computer programs for system performance effectiveness. Report N00140 666 0499, Research Triangle Institute, Triangle Park, North Carolina.
32. QUADE ES (1964) Analysis for military decisions. R-387-PR, RAND Corporation, Santa Monica, California.
33. PREST AR and TURVEY R (1966) Theories of cost benefit analysis. In *Surveys of Economic Theory*, Vol III. MacMillan, London.
34. ROSEN S (1964) Programming systems and languages, a historical survey. In *AFIPS Conference Proceedings*, Vol. 25. Cleaver-Hume, London.
35. SOLOMON MB, JR (1966) Economics of scale and the IBM/360. In *Communications of the ACM*.
36. STEDRY AC (1960) *Budget Control and Cost Behavior*. Prentice-Hall, New Jersey.
37. TUKEY JW (1962) The future of data analysis. *Ann. Math. Stat.* 33(1), 1–67.
38. TUKEY JW (1970) *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.
39. VAN HORN RL (1961) Systematic Methods for Programming Simplification. P-2447, RAND Corporation, Santa Monica, California.
40. VICKERY WS (1964) *Microstatics*. Harcourt, Brace & World, New York.
41. WEINBERGER AH (1963) Economic Evaluation of R&D Projects. *J. Chem. Engng* 17(5 & 6) 18(1–4).

ADDRESS FOR CORRESPONDENCE: *Donald V Mathusz Esq.*, 12726, *Ben Fry Drive*, *Chester*, *Virginia* 23831, *USA*.

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

Memoranda

Reducing the Number of Probabilistic Variables in Risk Simulation

THE VARIABLES in a risk simulation study can be thought of as falling into two categories. The first category (which will be referred to as category A) consists of those variables which, for the purposes of the analysis being undertaken, need to be described by a probability distribution; the second category (which will be referred to as category B) consists of those variables which can adequately be described by point estimates. This note summarises the results of some research carried out by the author into the extent to which a sensitivity analysis enables category A variables to be distinguished from category B variables.

Suppose that there are n variables and that:

$$P = f(X_1, X_2 \ldots . X_n)$$

where P is the value of the performance measure and X_i is the value of the i-th variable. There are a number of different ways in which a sensitivity analysis can be carried out. Here it will be assumed that, in addition to a most likely estimate E_i for each variable i, management make an 'high estimate' U_i and a 'low estimate' L_i. The output from the sensitivity analysis will be assumed to be $\{|S_i|: 1 \leqslant i \leqslant n\}$, where

$$Si = f(X_1, X_2 \ldots X_{i-1}, U_i, X_{i+1} \ldots X_n) - f(X_1, X_2 \ldots X_{i-1}, L_i, X_{i+1} \ldots X_n).$$

$U_i - L_i$ will be referred to as the range of variable i and $|S_i|$ as its sensitivity coefficient. (The precise meanings of the terms 'high estimate' and 'low estimate' need not concern us. In order that the $|S_i|$ be directly comparable it is only necessary for management to be consistent in their use of the terms. If the high estimate is considered to have a ten per cent chance of being exceeded for one variable then this should also be true for other variables, etc.)

In the case of the linear model:

$$P = \sum_{i=1}^{n} a_i X_i \qquad (a_i\text{'s constant})$$

it is easy to see that

$$S_i = a_i (U_i - L_i)$$

and straightforward analysis shows that:

(a) the effect of an error in the mean of variable i on the mean of the performance measure is equal to the product of S_i and the error (the error being expressed as a proportion of the variable's range);

(b) the effect of a certain percentage error in the standard deviation of variable i on the standard deviation of the performance measure is approximately proportional to the product of S_i^2 and the percentage error;

(c) the effect of total dependence between variables i and j on the standard deviation of the performance measure is approximately proportional to the product of S_i and S_j.

In a linear model the sensitivity coefficients do therefore have a clear interpretation in terms of the importance of the means and standard deviations of, and the coefficients of correlation between, different variables. To test their usefulness in the case of non-linear models five risk simulation case studies (to be referred to as cases A, B, C, D and E) have been examined in depth. Case A is based on the example in [2]. Case B is based on [3]. Cases C and D are based on two unpublished studies, the first involving a proposed investment in an advertising campaign for a consumer product, the second involving

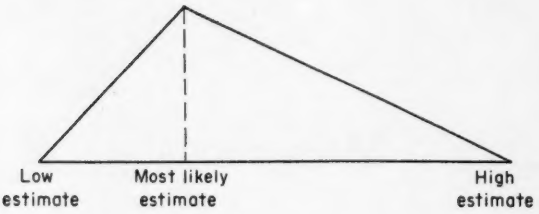


FIG. 1. *Triangular distribution used for category A variables.*

new plant for a major UK chemical company. Case E is based on [1].

For each variable in each case study a quantity R_i (to be referred to as the variable's range coefficient) was defined as follows:

$$R_i = \frac{|S_i|}{\text{Max } \{|S_k| : k = 1....n\}}$$

and the effect of using the rule:

If $R_i \leqslant C$ assign variable i to category B

If $R_i > C$ assign variable i to category A

was investigated for different values of C.

Four risk simulations (1000 runs each) were carried out for each of the five case studies. In the first of these simulations each of the variables was described by a triangular distribution (see Fig. 1). The other three simulations tested the consequences for C = 0·05, 0·1 and 0·25 of using the rule described above. Category A variables were described by the triangular distributions indicated in Fig. 1 and category B variables were described by their most likely estimates. The distributions which were used were in all cases assumed to be independent of each other and net present value (discount rate = 10%) was used throughout as the performance measure. The results obtained are summarised in Tables 1, 2 and 3 (note in the interpretation of tables 2 and 3 that C = O corresponds to the first simulation where all variables were described by distributions).

It can be seen from Table 3 that errors in the standard deviation of the performance measure resulting from the application of the rule described above are very small indeed even when C = 0·25. This is not surprising in view of the results mentioned above relating to the linear model which suggest that it is the magnitude of R_i^2 rather than R_i which determines the effect of the standard deviation of variable i on the standard deviation of the performance measure.

It can be seen from Table 2 that the errors in the mean of the performance measure resulting from the application of the rules which were used are on occasion quite large. Further investigations revealed that this is almost entirely because some category B variables were highly skewed and had the property that the difference between E_i and μ_i was quite large. Table 4 shows that the mean of the performance measure varies far less when category B variables are replaced

TABLE 1. NUMBER OF CATEGORY A VARIABLES FOR DIFFERENT VALUES OF THE 'CUT-OFF RANGE COEFFICIENT' C

Case study	Total No. of variables	Number of category A variables when C equals 0·05	0·1	0·25
A	9	7	5	4
B	12	10	6	4
C	7	7	7	6
D	17	8	7	6
E	4	4	4	2

Vol. 5
1977

TABLE 2. EFFECT ON MEAN OF NPV OF ASSUMING THAT, WHEN $R_i \leqslant C$ VARIABLE i IS EQUAL TO ITS BEST ESTIMATE

Case Study	Mean of NPV when C equals 0	0·05	0·1	0·25
A	−2590	−2470	−3020	−3030
B	1750	1860	2220	3060
C	74	74	74	78
D	8870	8870	8860	8880
E	2980	2908	2980	2630

TABLE 3. EFFECT ON SD OF NPV OF ASSUMING THAT, WHEN $R_i \leqslant C$ VARIABLE i IS EQUAL TO ITS BEST ESTIMATE

Case study	SD of NPV when C equals 0	0·05	0·1	0·25
A	10040	10030	9940	9940
B	5690	5670	5640	5670
C	93	93	93	94
D	535	533	532	524
E	1990	1990	1990	1920

TABLE 4. EFFECT ON MEAN OF NPV OF ASSUMING THAT WHEN $R_i \leqslant C$ VARIABLE i IS EQUAL TO ITS MEAN

Case study	Mean of NPV when C equals 0	0·05	0·1	0·25
A	−2590	−2580	−2630	−2650
B	1750	1740	1770	1770
C	74	74	74	75
D	8870	8870	8870	8870
E	2980	2980	2980	2950

by the mean of their distribution instead of by their most likely value.

As a general rule it seems sensible to set a category B variable equal to an estimate of μ_i rather than to E_i. The PERT formula

$$\mu_i = \tfrac{1}{6}[L_i + 4E_i + U_i]$$

or a modification of it could be used to provide the estimate of μ_i.

Further simulations were carried out to test the extent to which the three results (see a, b and c above) produced for a linear model hold for the non-linear model in case A. All three results were found to be approximately true [the results in (a) being a somewhat closer approximation to the truth than the results in (b) and (c)]. However, two features of a non-linear model should be borne in mind. These are:

(i) An error in the mean of a variable is liable to affect the standard deviation of the performance measure.

(ii) A dependence between two variables is liable to affect the mean of the performance measure.

Sensitivity coefficients provide no real indication as to the extent of these effects. If the effects are significant the results from a

sensitivity analysis are liable to be totally misleading. (The reader is invited to consider the usefulness of the output from a sensitivity analysis when $P = X_1X_2 + X_3$ and $L_1 = -11$, $E_1 = 1$, $U_1 = 13$, $L_2 = -12$, $E_2 = 0$, $U_2 = 12$, $L_3 = 0$, $E_3 = 12$, $U_3 = 24$).

To summarise: the use of a sensitivity analysis to identify category A and category B variables would have worked well in the five case studies which have been considered providing that the point estimates chosen for the category B variables were estimates of their means rather than 'most likely' estimates. However, a sensitivity analysis can in theory be totally misleading and the analyst should be aware of potential pitfalls.

REFERENCES

1. ECONOMOS AM (1968) A financial simulation for risk analysis of a proposed subsidiary. *Mgmt Sci.* **15**, B675–682.
2. HERTZ DB (1964) Risk analysis in capital investment. *Harv. Bus. Rev.* **42** (1), 96–108.
3. KRYZANOWSKI L, LUSTIG P and SCHWAB B (1972) Monte-Carlo simulation and capital expenditure decisions—a case study. *Eng. Econ.* **18**, 31–471

Dr JC Hull
(*March 1977*)
Cranfield School of Management
Cranfield
Bedford
England MK 43 0AL

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

Machine Scheduling with Mixed Integer Goal Programming

THE n-job, one-machine scheduling problem can be attacked using a mixed integer goal programming model. The integer variables serve to establish a sequence and to prohibit machine interference. The other variables, including the deviations from goals, are continuous. Such a model includes sequence constraints, noninterference constraints, and goal constraints. Two difficulties of previous efforts in machine scheduling are overcome; (1) mixed integer programming codes of sufficient capability are currently available, and (2) the model permits simultaneous consideration of multiple decision criteria.

The sequence constraints account for each possible permutation of n jobs which may be considered. The approach taken here is to consider the n jobs in a Cartesian product of ordered pairs using the linear order relation of 'immediate predecessor' to establish integer variables of the form:

$$u_{ij} = \begin{cases} 1 \text{ if job } i \text{ immediately precedes job } j, \\ 0 \text{ otherwise.} \end{cases}$$

The first job will have no predecessor and the last job will have no successor; therefore, integer variables of the form:

$$v_j = \begin{cases} 1 \text{ if job } j \text{ has no predecessor,} \\ 0 \text{ if otherwise} \end{cases}$$

and

$$w_i = \begin{cases} 1 \text{ if } i \text{ has no successor,} \\ 0 \text{ otherwise} \end{cases}$$

are established. The sequencing constraints are then written as:

$$\sum_{i \neq j} u_{ij} + v_j = 1 \qquad j = 1, \ldots, n,$$

$$\sum_{j \neq i} u_{ij} + w_i = 1 \qquad i = 1, \ldots, n,$$

$$\sum_{i=1}^{n} v_j = 1 \quad \text{and} \quad \sum_{i=1}^{n} w_i = 1$$

and where $\sum_{i \neq j}$ is equivalent to $\sum_{\substack{i=1 \\ i \neq j}}^{n}$, etc.

Non-interference constraints are needed to insure that jobs do not overlap in processing. An additional variable needed for these constraints is t_i, the start time for job i. Other additional parameters are:

M, a large constant $\gg \max \{t_i\}$;
P_i, run time for job i;
m_{ij}, makeready (setup) for job j when preceded by i;
b_i, makeready when job i has no predecessor.

The non-interference constraints are:

$$(t_j - t_i) + M(1 - u_{ij}) \geq p_i + \sum_{k \neq i} m_{ki} u_{ki} + b_i v_i,$$

for all i and j, $i \neq j$.

The goals selected for use in this problem are:

(1) minimize makeready (sequence dependent),
(2) minimize tardiness of jobs,
(3) minimize finished inventory (earliness).

Constraints for the makeready goal are:

$$\sum_{i \neq j} m_{ij} u_{ij} + b_j v_j - d_j \leq m_j$$

for $j = 1,\ldots,m$, where $m_j = \min\{m_{ij}, i=1,\ldots, n, i \neq j\}$ and d_j is the overachievement of the goal for job j. Since the tardiness and earliness goals are the overachievement and underachievement of a desired finish time respectively, they are included in the same set of constraints:

$$b_j v_j + \sum_{i \neq j} m_{ij} u_{ij} + t_j - e_{j1} + e_{j2} = f_j - p_j \text{ for } j = 1, \ldots, n,$$

where e_{j1} is the overachievement (lateness) for job j,
e_{j2} is the underachievement (earliness) for job j,
and f_j is the desired finish time for job j.

In an actual press environment in Phoenix, sets of jobs to be analyzed by the proposed model generally range in size from 2 to 10. To illustrate the model, a 3-job situation from this environment at W. A. Krueger Company was analyzed. The model was solved by using a mixed integer programming computer code (MINT) taken from Kuester and Mize.

Both the tractability of the model and its practicability are of interest. The size of the problem formulation generated, particularly the number of integer variables, determines the feasibility of applying a mixed integer programming code. In general, for this type formulation, the numbers of variables and constraints are: $2n + n(n-1)$ integer variables (0–1), $4n$ real variables, $n(n-1) + 2n$ ($\leq$) constraints, and $3n + 2$ ($=$) constraints, where n is the number of jobs being scheduled. For the 3-job, 3-objective, one-machine example that was solved using historical data, 24 variables, and 20 constraints were needed to find a solution that could be implemented at the Phoenix plant.

The range of 2–10 jobs indicated for the actual environment studied thus yields problems with reasonable size for solution on large scale computers. Approximating the larger situations by breakdown into sets of about 6 jobs would yield reasonable results at W. A. Krueger Company since delivery requirements force relatively little interaction among jobs at the beginning and ending of the larger job set.

The other concern is practicability. First, the formulation must generate different solutions than normally generated by subjective managerial analysis. For the example 3-job problem, the objective function based on managerial input is:

$$\sum_{j=1}^{n} (10\, e_{j1} + 5\, d_j + 1\, e_{j2}).$$

Using this objective function, the three jobs should be ordered 2–1–3 based on the mixed IP solution; actual implemented order was 1–2–3. The data used are available on standard documents and thus are not expensive to obtain. Frequency of application should be about once every 2 weeks since 2–4 jobs are generally processed per week.

To enhance tractability, some of the more efficient algorithm codes using 0–1 integer variables can be applied to the model as formulated. Practicability can be enhanced by developing a pre-code program to accept raw data and generate the standardized formulation.

REFERENCE

1. Kuester JL and Mize JH (1973) *Optimization Techniques with FORTRAN.* McGraw-Hill, New York.

JC Hershauer
JW Gowens
(*March 1977*)

College of Business
Division of Management
University of Texas at San Antonio
San Antonio
TX 78285
USA

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

A Framework for the Analysis of the cost of Police Patrol Vehicle Breakdowns

INTRODUCTION

ONE CONCERN of police officials is the decision relative to the type and level of resources to be allocated to different patrol areas. Regardless of the resource allocation decision, any factors which depreciate the planned resource levels within the patrol areas can be expected to lower the effectiveness of the patrol activities. This paper considers one such factor: the breakdown of patrol vehicles and the implicit cost of the breakdowns. If a police patrol vehicle must be removed from service in order to repair mechanical defects, the quality (and effectiveness) of police services will not be at the desired (planned) level. Given the high cost of fielding a police patrol car around-the-clock (one estimate suggests the annual cost for a two man vehicle is from $100,000 to over $200,000 [2:3]), the time lost due to mechanical failure can be very costly.

This memorandum examines a framework for the determination of the implicit cost of time lost from police patrol due to police vehicle breakdowns, and the resulting implications for management decisions. The result of an actual study in which the approach was implemented yields cost of a magnitude that indicate investment in a preventive maintenance program may be worthwhile; furthermore, policy changes may result in substantial savings.

A FRAMEWORK FOR THE ANALYSIS

The framework is based upon a simulation methodology which proved successful in the analysis of an actual police department. The general framework will be discussed in the context of the specific case in which the approach was used. However, it may be expected that some of the characteristics of this case may suggest some general principles to be used in the analysis of the cost of police patrol car breakdowns in other cities.

The study took place in a city of approximately 175,000 persons. Three-quarters of the $400,000 police department budget is devoted to personnel costs, supporting approximately 250 personnel. There were approximately 87,000 calls for service during the year of the study. The city is divided into twelve patrol zones and the patrol operations consist of three shifts: day, evening, and night. When sufficient personnel and cars are available, one car is assigned to each zone. When insufficient cars are available for all zones, less active zones are combined in pairs to be patrolled by a single vehicle. Patrol cars return to headquarters at shift change and the vehicles are turned over to the next shift; thus the cars are in use 24 hr a day.

The 'round-the-clock' use of the cars, in conjunction with the lack of a preventive maintenance program, results in the breakdown of vehicles while in service. (In this note, 'breakdown' refers to any need for service; thus a burned out headlight, a flat tire, and an overheating engine would be considered to be breakdowns). The study was undertaken in order to determine the implicit cost of the vehicle breakdowns: the cost of the lost police service resulting from the breakdowns. (There is another cost related to the breakdowns—increased response time which will not be considered here.)

A simulation model of the breakdown process was developed in order to analyze the problem. Thus, it was necessary to collect data concerning breakdowns on each shift from the duty logs of the individual police officers. The data suggest that two types of breakdowns exist: regular breakdowns and special breakdowns. The duty logs indicate that, in a majority of the cases, the officers were out of service at the maintenance facility for less than 15 min. These breakdowns are defined as regular breakdowns, and it may be assumed that many of these breakdowns are for the purpose of refueling.

Special breakdowns, on the other hand, are of longer duration, the majority ranging from 20 to 45 min. The maximum recorded maintenance time was 80 min. The breakdown data were used to construct both regular breakdown and special breakdown distributions for each shift. These distributions provide the probability that a given number of cars will require maintenance on a particular shift. The data also provide information on the distribution of maintenance time for the special breakdowns on each shift. The simulation model is based upon these empirical distributions. (It may be shown that the Poisson is a good model of the breakdown process, and an analysis of such a model, including goodness-of-fit tests and implications of the Poisson for maintenance facility staffing, appears in [1]).

The model simulates the breakdown of vehicles on each of the three shifts for a 365-day period. A brief review of the simulation methodology for each shift follows: the number of regular and special breakdowns for the shift is determined by generating a random number from the regular and special breakdown distributions. Since cars are randomly assigned, it is assumed that vehicle breakdowns are randomly distributed over the twelve patrol areas. Thus, the particular patrol area affected by a given breakdown depends upon the value of a random number generated from a uniform distribution. The number of police officers in the vehicle is determined by drawing a number from a distribution which embodies the staffing policy for the shifts. Lower and upper limits for travel time from patrol beats to the maintenance area were defined by police personnel; in the simulation, travel time from a given patrol beat to the maintenance area is based upon a uniform distribution with lower and upper limits as specified by the police administrators. Maintenance time for regular breakdowns is uniformly distributed over the interval 10–15 min. Maintenance time for special breakdowns is determined by generating a number from the maintenance time distribution for special breakdowns. It is assumed that the vehicle will return by the most direct route to the patrol area, thus the return time to a given area is a constant: the lower limit on travel time defined for that area. Travel, maintenance and return times are multiplied by the number of officers in the car in order to determine the lost time for the respective categories.

RESULTS AND IMPLICATIONS

The magnitude of the implicit costs resulting from vehicle breakdowns is exemplified by the results of the case considered herein. In this section we examine these costs, and the resulting implications for management policies and decisions. The system was simulated under several different assumptions and for several different policies. This discussion will focus first on the actual costs of lost police officer patrol time resulting from the vehicle breakdowns. An analysis of the results from another point of view—lost patrol vehicle time—will be discussed at the end of this section.

The total lost time—the sum of travel time, maintenance time, and return time—during the 365-day period is equivalent to the loss of 5·19 police officers annually. At an average annual cost in excess of $13,000 (salary and direct expenses) per police officer, the cost of the lost time is $67,470.

The large number of regular breakdowns reported in the duty logs suggests that perhaps many of these breakdowns are for purposes of refueling. Since refueling is a necessity, it may be desirable to determine the lost time in excess of time lost due to refueling; i.e. assuming that refueling time is a sunk cost, how much lost time can be attributed to factors other than refueling? In order to determine this amount three different assumptions are made concerning refueling, namely that 25%, 50%, and 75% of the regular breakdowns are for the purpose of refueling. The results of the simulation under these assumptions are that the lost time is equivalent to 4·73, 4·27, and 3·80 police officers annually respectively at the annual costs of $61,490, $55,510, and $49,400. It should be noted that these costs are based on all lost time except for the actual time spent in refueling—lost travel time and return time necessitated by the need for refueling are included in the computations, since, as noted below, it may be possible to eliminate these times by a policy change.

Since many of the regular breakdowns are for purposes of refueling, it can be expected that establishing a policy relative to refueling will result in less lost time and a reduction in the implicit cost of the lost time. Since the cars from one shift are driven to police headquarters where they are turned over to the next shift, a policy could be established whereby one-third of the cars are refueled at each shift change. Under this policy, the time

Vol. 5
1977

lost for refueling will be the same as it is under the present system, but the travel time to and the return time from the maintenance facility will be eliminated under this policy. In order to determine the resulting savings, the system was simulated under the assumptions that 25%, 50% and 75% of the regular breakdowns are for purposes of refueling. The policy change savings are \$9,230, \$18,590 and \$27,690, respectively, in comparison to the cost of the present system under the respective assumptions.

The preceding analysis considered the costs of patrol time lost by the police officers affected by the breakdown. Another point of view is that the cost of the lost patrol vehicle time should be examined. That is, the breakdown removes the patrol car from service for a specific period of time. Given the annual cost of fielding a patrol car, what are the implicit costs of this lost vehicle time? In these simulations the travel time, maintenance time, and return time are not multiplied by the number of police officers in the affected car—we were concerned with lost patrol vehicle time, not lost police officer time. Total time lost during the 365-day period is equivalent to the loss of 0·72 'round-the-clock' police patrol vehicles. To fill one post 24 hr per day, 365 days per year requires approximately five individuals (this takes into account vacations, days off, and sick leave) [2]. Thus a two person vehicle requires approximately ten individuals. Since annual salary and direct costs exceed \$13,000 per police officer, the implicit cost of the lost time for 0·72 vehicles is \$93,600. If 25% of the regular breakdowns are for purposes of refueling, the total lost time is equivalent to 0·66 'round-the-clock' patrol vehicles, for an implicit annual cost of \$85,800. The 50% and 75% assumptions yield 0·59 and 0·53 lost vehicles annually at respective costs of \$76,700 and \$68,900. Policy change savings under the 25%, 50%, and 75% assumptions are \$13,000, \$24,700, and \$39,000, respectively. It should be noted that these analyses are based upon two-person vehicles and that all of the patrol vehicles are not two-person vehicles. Thus these figures can be considered to be upper limits on the implicit costs resulting from the police vehicle breakdowns,

The results of the simulation under the various assumptions and for the various policies provide estimates of the costs of lost time as a result of vehicle breakdowns. The costs range from \$49,400 to \$93,600 with policy change savings ranging from \$9,230 to \$39,000. The magnitude of these costs suggests that it may be worthwhile to establish a preventive maintenance program for the police vehicles. The costs of a preventive maintenance program should be evaluated in light of the implicit costs resulting from the loss of police services due to patrol car breakdowns. Furthermore, it appears that simple policy changes may result in substantial implicit savings.

REFERENCES

1. GLEASON JM A Poisson model of a police patrol car breakdown process: implications for maintenance facility staffing. Institute for Applied Operations Research in Public Administration, College of Business Administration, Texas Tech University, Lubbock.
2. LARSON RC (1972) *Urban Police Patrol Analysis*, MIT Press, Cambridge.

JM Gleason
(*March* 1977)
College of Business Administration
University of Nebraska at Omaha
Omaha
Nebraska 68101
USA

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

Logic Tree Analysis for Decision Making

INTRODUCTION

THIS ARTICLE describes and illustrates an improvement to the traditional approach of using Decision Tree Analysis [1–3]. This approach has been titled Logic Tree Analysis due to the nature of the problem formulation and its solution. A decision tree is actually just a subset of a logic tree which contains only 'OR' logic decision and chance branching. The motivation for this approach is based on the observation that by adding 'AND' logic elements to the standard decision tree format it might be possible for more complex decision situations to be analyzed. Thus, it might be an advantage to consider two or more simultaneous choices of ventures entered into simultaneously. The standard decision tree approach permits evaluation and ranking only of individual outcomes, that is those outcomes not involving more than one venture–outcome combinations at a time.

Logic tree analysis employs both 'AND' and 'OR' (decision or chance branching) gates where combinations of more than one event are of importance. This analysis is formulated in such a manner to describe how a given type of outcome can occur. These outcomes usually include the simultaneous occurrence of more than one event and for technical systems of interest can easily number hundreds or thousands of outcome sets which include a few events.

APPLICATION

First it should be clearly stated that there can be no advantage to analyzing combinations of mutually independent alternatives with logic tree analysis which is precisely equivalent to performing a decision tree analysis. The alternatives might just as well be analyzed by decision tree analysis. However, if several ventures which are considered simultaneously do interact in some definable manner, and if that dependency can be shown and calculated as a function of the composition of the members of an identified set, then there could be an advantage in using logic tree analysis to do the analysis.

For purposes of illustration, suppose that a company is considering producing an indeterminate number of new products at a given location. A new building, of presently indeterminate size, will have to be constructed and each product will be assigned its allocated share of the facility depreciation, overhead costs and other indirect costs. The share of the total cost for each product will depend on how many products are produced and which ones, because that will define the total costs. Also to be considered are the likely market size and probability, sales price, and direct manufacturing cost for each product exactly as used by other risk analysis techniques. As another example, suppose that several possible products have the probability of success dependent on the success of one or more other products. The sales probability of the second series of products depends on how successfully the first products are marketed. In both the examples discussed, the probability of success for combinations of ventures is different for the ventures considered separately than for combinations of them. Of course, the logic tree approach does not decrease the importance of having valid probability data. It is further complicated because in addition to independent probabilities, now dependent probabilities and contingent costs are also required.

EXAMPLE

The following example will be illustrated to show how the full logic tree approach might be used by a company. The decision for this particular problem is formulated as "How the company can make revenue from a Product Mix". The information for this example is shown on Table 1.

As can be seen from Table 1, seven different products (A–G) can be selected for market. Each of these seven products has a different initial investment, a different estimate of the

Vol. 5
1977

TABLE 1. PRODUCT CHARACTERISTICS

Product	Initial investment	Demand	Corresponding probability	Units / Year	Direct profit $/unit	Space required (ft²)
A	$15,000	Small	0·20	20,000	$2·20	10,000
		Moderate	0·45	40,000		
		Large	0·35	70,000		
B	$11,500	Small	0·35	8000	$14·45	28,000
		Moderate	0·40	20,000		
		Large	0·25	40,000		
C	$25,000	Small	0·50	17,000	$3·31	21,000
		Moderate	0·35	40,000		
		Large	0·15	85,000		
D	$5500	Small	0·43	800	$14·07	4500
		Moderate	0·37	1500		
		Large	0·20	3000		
E	$87,000	Small	0·33	50,000	$2·18	42,000
		Moderate	0·58	200,000		
		Large	0·09	650,000		
F	$8900	Small	0·52	400	$50·27	9800
		Moderate	0·33	900		
		Large	0·25	2000		
G	$16,470	Small	0·11	1100	$11·85	23,500
		Moderate	0·74	8750		
		Large	0·15	26,420		

TABLE 2. ABBREVIATED RESULTS

Product(s)	$ All small	$ All moderate	$ All large	$ Cumulative
A	−102,304	−152,160	− 27,318	−281,782
B	−128,707	122,229	363,605	357,127
B, C	−497,303	−112,496	58,011	−551,788
B, D	−511,462	−212,814	−54,144	−778,920
A, E, F, G	−591,503	460,326	226,609	95,433
B, E, F, G	−529,241	709,159	617,532	797,451
A, B, C, D, E, F, G	−496,251	954,540	959,901	2,600,382

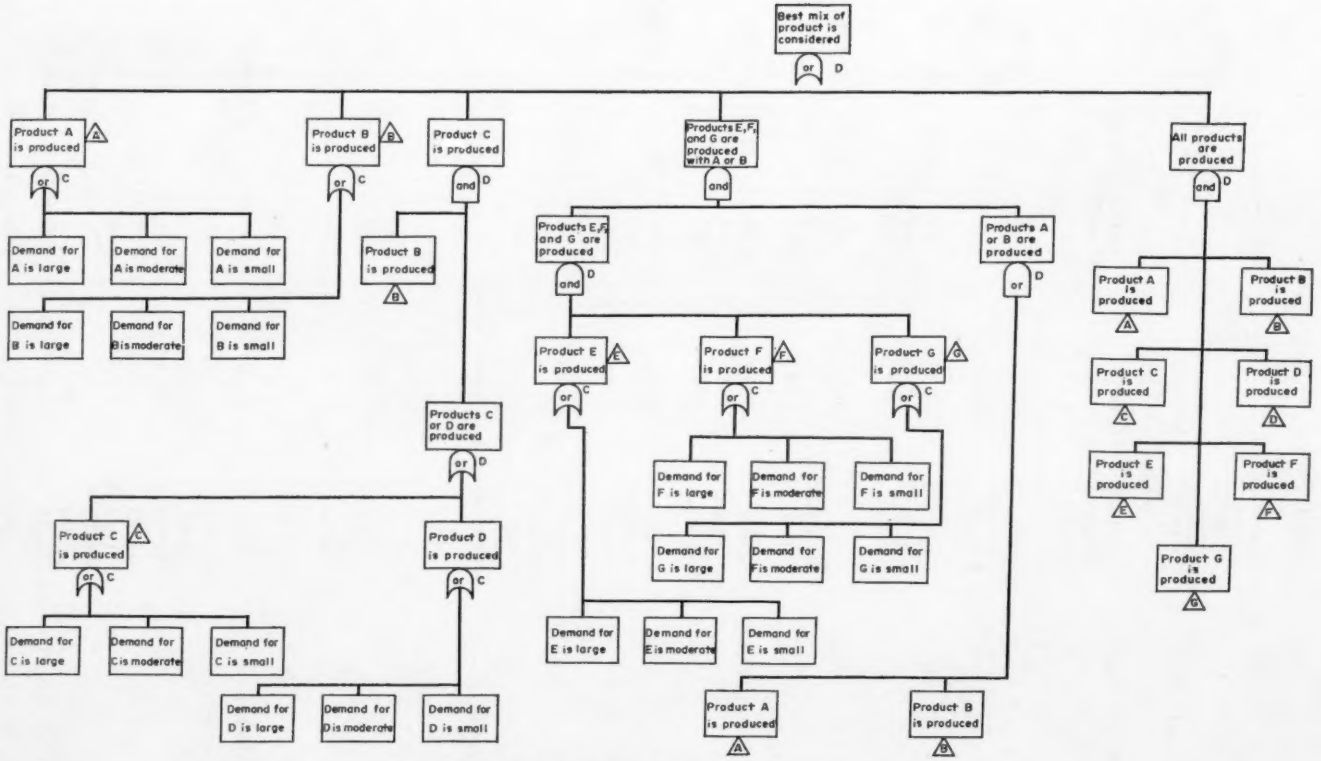


FIG. 1. *Logic tree diagram (success tree)*

potential demand, a different direct profit per unit, and a different space required for the manufacture of the product involved. The logic tree does not show those paths which do not produce revenue. Also, no attention has been given to actions of competitors in this example, although such a factor could be incorporated into the analysis. Figure 1 depicts a graphical layout of the decision process. Standard logic tree symbols, similar to electrical engineering logic gate symbols, are used in the tree diagram instead of decision tree symbols, but a 'D' for decision choice or a 'C' for chance event has been placed by each gate symbol to identify how the branching occurs. Also, the diagram proceeds from top to bottom in standard logic tree layout rather than from left to right.

The logic tree in Fig. 1 is actually a success tree which details management's constraints and preferences for a proposed marketing mix. The firm might consider marketing A only or B only, or B with C or D, or E, F, and G with A or B, or A, B, C, D, E, F, and G simultaneously. The total cost to build plus overhead is represented by $150,000 plus $2·00 per square foot of space needed. The products manufactured require space independent of the market demand which develops and will share in the cost as a direct proportion of the space required. Each product also requires an investment for manufacturing equipment and it has been determined that each product has discrete probabilities of having high, moderate, or low market demands. The life of the project is expected to be 5 years, and the market rate of interest is 8·5 per cent.

Table 2 provides the results of the analysis for those cases with sales corresponding to all market demands less facility overhead cost. The discounted cash flow returns were computed and multiplied by the individual fraction probability of achieving the indicated level of sales. For the last case, only three from a possible 2187 permutations were evaluated.

SUMMARY

It is clear that the full logic tree approach can be a useful tool for evaluating interdependent ventures where the number of permutations possible is very large. Logic tree analysis computer programs already exist [4]. These computer programs usually find the sets and also typically perform probability calculations. Previous programs, now considered to be obsolete, used Monte Carlo calculations to find the sets and could therefore not separate the quantification characteristic (probability) from the task of finding the set. Reference 4 details one such program intended to be used for Fault Tree Analysis.

REFERENCES

1. HAMMOND JS (1967) Better decision with preference theory. *Harv. Bus. Rev.* **45,** 123.
2. MAGEE JF (1964) Decision trees for decision making. *Harv. Bus. Rev.* **42,** 126.
3. MAGEE JF (1964) How to use decision trees in capital investment. *Harv. Bus. Rev.* **42,** 79.
4. VAN SLYKE WJ (1975) Allcuts, a fast, comprehensive fault tree analysis code. ARH-ST-112, Atlantic Richfield Hanford Company.

David A Lopez
William J Van Slyke
(*May* 1977)

Graduate School of Business Administration
University of Washington
Seattle
Washington 98195
USA

OMEGA, The Int. Jl. of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain.

Urban Recreational Planning: A Proposed Methodology

ONE OF the responsibilities of urban planners who function within municipal governments is to provide appropriate recreational facilities using limited resources for the residents of the urban area. However, this is a problem that is more complex than the traditional capital budgeting problem for several reasons. First, it is difficult to measure the return on a project which has a social function as the primary consideration for its undertaking. Second, the existence of various social conditions (i.e. economic and racial segregation) creates conflicting demands and objectives for recreational facilities. In addition, the normal constraints of limited funds both for construction and maintenance and limited land resources further complicate the selection process.

While the recreational facility selection problem has received little attention specifically, there have been a number of quantitative methodologies applied to the general problem of capital project selection in urban planning. These approaches include rank-ordered decision making, program budgeting, and benefit–cost analysis. All of these techniques attempt to incorporate some measure of goal achievement into the evaluation process by assigning a numerical or dollar value measuring goal effectiveness. However, such assignments are always somewhat arbitrary by their very nature and thus subject to suspicion. Alternatively, mathematical programming has been employed on several occasions for urban planning [3, 5]. However, in these applications, an attempt is made to translate incommensurable goals for urban projects into monetary units and then combine them into a unidimensional objective function. This is an objectionable procedure as such translations lead to an arbitrary fabrication of the problem.

An alternative to these approaches is integer goal programming. Goal programming lends itself to problems with incommensurable goals since it employs a multidimensional objective function. The traditional goal programming model is combined with a 0–1 integer programming model [1] to allow for the indivisibility of projects and avoid the problem of suboptimal results [5].

In order to apply the integer goal programming model, the urban area under analysis should be subdivided into sectors with common economic and racial characteristics. Recognizing the fact that different income and racial groups have different interests in terms of recreation type and different economic means to support recreation, alternative facilities should be identified for each sector [2, 4]. The most common recreational facilities for urban areas are swimming pools, tennis courts, gymnasiums, athletic fields, and golf courses. As such, the variables of the proposed model would identify specific facilities within the alternative sectors. Financial goal-constraints of the model reflect land cost, construction cost, private gift and public grant contributions, and operating budget. An additional resource constraint is the land available for construction.

Goals are established for expected user-days per facility per elapsed time period. Subgoals should also be established for expected user-days for both low income and high income and racial use. Goal levels are established for equal facility use for the different groups so that specific areas are not neglected because of income or racial characteristics. Additional goal-constraints could reflect the demand for specific facilities in certain sectors and the needs of special groups (such as the elderly).

The objective function is developed according to their perceived importance as ascertained by urban officials and residents. Typically, resource goals must be given high priority but not however at the expense of equal facility use in low income and minority areas. The integer goal programming model enables the planner to experiment via sensitivity analysis with alternative goal levels and priority structures.

REFERENCES

1. LEE SM (1976) *Linear Optimization for*

Vol. 5
1977

Management. pp. 421–446. Petrocelli/Charter, New York.
2. RICHARDSON HW (1973) A comment on some uses of mathematical models in urban economics. *Urb. Studies* **10**, 259–270.
3. SCHLAGER KJ (1965) A land-use plan design model. *J. Am. Inst. Plan.* **31**, 103–117.
4. SHILTZ T and MOFFITT W (1971) Inner-city/outer-city relationships in metropolitan areas. *Urb. Affairs Q.* **7**, 75–108.
5. STUART DG (1972) Urban improvement programming models, in *Decision Making in Urban Planning* (Ed ROBINSON IM) pp. 342–377. Sage Publications, Beverly Hills, CA.
6. WEINGARTNER HM (1963) *Mathematical Programming and the Analysis of Capital Budgeting Problems*. Prentice-Hall, Englewood Cliffs, NJ.

Bernard W Taylor, III
Arthur J Keown
(*April 1977*)
Department of Business Administration
College of Business
Virginia Polytechnic Institute and State University
Blacksburg
VA 24061
USA

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 5, 1977. Pergamon Press. Printed in Great Britain

International Register of Theses and Dissertations

ALCANTARA B (1972) *French: Optimisation et Simulation en Avenir Aléatoire—Applications a la Gestion des Stocks.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

BARIANI R (1972) *French: Etude de la Politique des Dividendes des Grandes Sociétés Françaises.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

BENOIT R (1974) *French: Fonds d'investissement Régional.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

DERYMCK J (1971) Simulation of the Launching of Industrial Products. MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

HEYDENREICH HD (1973) *German: Problembereiche ausländischer privater Direktinvestitionen—Eine empirische Untersuchung.* PhD Thesis, George-August-Universität, Göttingen. (Available at: Niedersächsischen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prinzenstrasse 1, Germany)

JENSSEN J (1973) *German: Die Finanzierung von Investitionen aus Abschreibungen in einer gleichgewichtigen Wirtschaft.* PhD Thesis, George-August-Universität, Göttingen. (Available at: Niedersächsischen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prinzenstrasse 1, Germany)

KAHLE E (1973) *German: Der Beitrag von Zielen und Entscheidungsmodellen zur Losung betriebswirtschaftlicher Probleme.* PhD Thesis, George-August-Universität, Göttingen. (Available at: Niedersächsischen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prinzenstrasse 1, Germany)

LEBLANC M (1973) *French: Développement des Systèmes d'Information—Centralisation—Decentralisation.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

LEJAY J (1973) *French: Modèle de Planification pour une Entreprise de Bâtiment et de Travaux Publics.* MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

MEINZEN F (1973) *German: Gibt es eine optimale Grösse von Elektrizitasverteilungsunternehmen? Ein Beitrag zur Losung der 'Strukturprobleme' auf dem Absatzmarkt für elektrische Energie.* PhD Thesis, George-August-Universität, Göttingen. (Available at: Niedersächsischen Staats-und Universitäts-Bibliothek, 240 Göttingen, Prinzenstrasse 1, Germany)

OHNISHI T (1973) *German: Tarifpolitik Preussens bis zur Gründung des Deutschen Zollvereins.* PhD Thesis, George-August-Universität, Göttingen. (Available at Niedersächsischen Staats-und Universitäts-Bibliothek. 2400 Göttingen, Prinzenstrasse 1, Germany)

PLATH F (1973) *German: Ein Beitrag zur ökonomischen Bewertung von Planungsvarianten zum Ausbau stadtischer Verkehrssysteme.* PhD Thesis, George-August-Universität, Göttingen. (Available at: Niedersächsischen Staats-und Universitäts-Bibliothek, 2400 Göttingen, Prinzenstrasse 1, Germany)

VU QH (1972) *French: Similarité Psychologique et non-Disponibilité comme Variables de la*

Substitution de Marques. MSc Thesis, Université de Sherbrooke. (Available at: the General Library, Université de Sherbrooke, Québec J1K 2R1, Canada)

The following are obtainable at the Joseph Regenstein Library, University of Chicago, 1100 East 57th Street, IL 60637, USA:

BELL PC (1976) A Finite Horizon Capital Installation Cost Problem. PhD Thesis, University of Chicago.

BLOCK MW, JR (1974) Member Commitment to Group Decisions: A Study of Individual and Group Determinants. PhD Thesis, University of Chicago.

CHELIUS JR (1974) An Analysis of Industrial Safety Regulation. PhD Thesis, University of Chicago.

GALAI D (1975) Pricing of Options and the Efficiency of the Chicago Board Options Exchange. PhD Thesis, University of Chicago.

IBBOTSON RG (1974) Price Performances of Common Stock New Issues. PhD Thesis, University of Chicago.

KRIDER CE (1976) Patterns and Determinants of Strikes in Non-Education Local Government. PhD Thesis, University of Chicago.

LEE MH (1974) Excess Inflation and Currency Depreciation. PhD Thesis, University of Chicago.

LEVICH RM (1976) The International Money Market: Tests of Forecasting Models and Market Efficiency. PhD Thesis, University of Chicago.

MCGUIRE CC, JR (1974) The Depreciation of Housing. PhD Thesis, University of Chicago.

MORGAN LG (1974) Stock Prices and Heteroscedasticity. PhD Thesis, University of Chicago.

NOWELL G (1976) A Study of Black Students and Their Experiences in the 'Careers for Blacks in Management Program'. PhD Thesis, University of Chicago.

PEACOCK P (1975) Defining Market Segments with Multidimensional Purchase Behavior. PhD Thesis, University of Chicago.

SCHLEEF HJ (1976) Optimal Control Models for Multi-Server Exponential Queueing Systems. PhD Thesis, University of Chicago.

SPENCE RJ (1974) Exclusive Dealing—Its Rationale and Effect on Competition: A Case Study. PhD Thesis, University of Chicago.

VALESCO JB (1976) Philippine Coconut Exports: A Study on the Feasibility of Stabilizing Export Earnings. PhD Thesis, University of Chicago.

The following are obtainable at the Library, Cranfield Institute of Technology, Cranfield, Bedford MK43 0AL, UK:

GLASSMAN DM (1972) The Community and Youth Service Association. MBA Thesis, Cranfield Institute of Technology.

GOLBEY A (1972) The Marketing of Contract Mining Skills. MBA Thesis, Cranfield Institute of Technology.

GOSLING MC (1972) The Design of a Management Control System for a New Factory. MBA Thesis, Cranfield Institute of Technology.

GRUBB RF (1971) A Vertical Profit Analysis and Market Study of Textured Tricel Yarns Used in Apparel End Uses. MBA Thesis, Cranfield Institute of Technology.

HERD DM (1972) A Study of Nuclear Fuel Leasing in Selected European Countries. MBA Thesis, Cranfield Institute of Technology.

HOOPER EJ (1970) A Survey of the Market for Industrial Flowmeters with Particular Reference to the Thodes Flowmeter. MBA Thesis, Cranfield Institute of Technology.

JEFFERIES RJS (1972) The Development of a Successful Airline. MBA Thesis, Cranfield Institute of Technology.

JOHNSON GL (1972) An Investment Appraisal of Selected Companies in the Chemical Sector of Industry. MBA Thesis, Cranfield Institute of Technology.

JONGEBREUR RS (1971) An Analysis of a Sub-Contract Machining Company. MBA Thesis, Cranfield Institute of Technology.

KENNEDY DP (1972) A Study of Some Aspects of Newspaper Marketing. MBA Thesis, Cranfield Institute of Technology.

MCGARVEY A (1972) A Market Investigation for a Manufacturer of Industrial Compressors. MBA Thesis, Cranfield Institute of Technology.

MILLER R (1930) Organisation at Brathay, MBA Thesis, Cranfield Institute of Technology.

PEARSON GWR (1972) Economic and Marketing Aspects of Air Freight, with Special Reference to Air Freight Forwarders. MBA Thesis, Cranfield Institute of Technology.

POWELL-SHEDDEN HRG (1970) The Evaluation of the Possibilities of Commercial Radio for xyz Press in a Given Area. MBA Thesis, Cranfield Institute of Technology.

RIQUEZES J (1971) Operational Research and the Social Sciences, with Special Reference to Urban and Regional Planning. MBA Thesis, Cranfield Institute of Technology.

SEMPER BM (1971) Labour Reporting—An Investigation of the Labour Reporting System at Mallory Batteries Limited. MBA Thesis, Cranfield Institute of Technology.

SIMPSON RAP (1972) Some Aspects of Management Control Systems within a Large Agricultural Company. MBA Thesis, Cranfield Institute of Technology.

SMITH MJ (1972) Operations Management in a Small Company. MBA Thesis, Cranfield Institute of Technology.

SONGHURST JV (1971) The Supply and Demand for Beef and Veal in the United Kingdom and the European Economic Community. MBA Thesis, Cranfield Institute of Technology.

THORNTON JGT (1972) The Warwick Pump and Engineering Company Limited—Introducing Long Range Planning. MBA Thesis, Cranfield Institute of Technology.

TOGEL ID (1972) The West European Toy Market—A Structural Study of the Market, its Suppliers and Environmental Conditions. MBA Thesis, Cranfield Institute of Technology.

TOMES F (1972) An Investigation with the Objective of Recommending how Ferodo's Present Manufacturing Capability Can Be Adapted to Meet Future Requirements for Asbestos Based Clutch Facings. MBA Thesis, Cranfield Institute of Technology.

TORVAS DP (1970) An Investigation and Techniques of Multi-Item Inventory Control. MBA Thesis, Cranfield Institute of Technology.

TSAROUCHIS NG (1971) Forecasting Time Series and their Simulation on Analogue Computer. MBA Thesis, Cranfield Institute of Technology.

USHERWOOD R (1970) Market Forecasting for Hot Finished Tube. MBA Thesis, Cranfield Institute of Technology.

VAFIADIS CG (1970) The Effect of Incomplete Information on Performance of Stock Control Schemes. MBA Thesis, Cranfield Institute of Technology.

VANDER MEERSCH R (1970) Technological Innovation, its Nature and Management. MBA Thesis, Cranfield Institute of Technology.

VARLEY A (1970) Report on Marketing Policies of BTR Polymer Fabrication Group—Technical Section. MBA Thesis, Cranfield Institute of Technology.

VLAHODIMOS C (1970) An OR Approach to the Potential Use of Hovercraft and the Hydrofoil in Restricted Waters—Application in the Aegean Sea. MBA Thesis, Cranfield Institute of Technology.

WALLER AG (1971) An Operational Research Analysis of a Continuous Weld Tube Plant (Corby Tube Works). MBA Thesis, Cranfield Institute of Technology.

WARD JG (1971) An Investigation of the Greek Tourist Industry, and an Evaluation of Market Potential for Thomson Holidays Ltd. MBA Thesis, Cranfield Institute of Technology.

WHITELEY N (1972) An Investigation into the Marketing Opportunities for Rainbow Trout. MBA Thesis, Cranfield Institute of Technology.

WHITTOW AF (1972) A Profile of the UK Horticultural Equipment Industry. MBA Thesis, Cranfield Institute of Technology.

WHITTOW J (1972) Ford of Europe and Business Graduates. MBA Thesis, Cranfield Institute of Technology.

WILLANS JRD (1970) The Opportunities of Civil Engineering in South Africa. MBA Thesis, Cranfield Institute of Technology.

WOOD KN (1972) Wage Differentials—A Pilot Study for the National Economic Development Office. MBA Thesis, Cranfield Institute of Technology.

WRIGHT RC (1971) A Comparison Between Various Optimization Techniques Used in Computer Simulation. MBA Thesis, Cranfield Institute of Technology.

YOUNG ARA (1970) An Investigation of Price, Cost, and Profitability Relationship in Aero Engine Production. MBA Thesis, Cranfield Institute of Technology.

The following are obtainable at the MIT Microreproduction Laboratory, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02139, USA:

BEARD GW (1973) Labor Versus the Multinational Corporation. SM Thesis, MIT, Cambridge, MA.

BLOUNT TD (1973) Design of a Man–Machine Decision System for a Production Scheduling Problem. SM Thesis, MIT, Cambridge, MA.

CHESNEY RA (1973) The Impact of Social Issues in the Banking Industry. SM Thesis, MIT, Cambridge, MA.

CLIFT LM (1973) The Management Function in Public Higher Education in Massachusetts. SM Thesis, MIT, Cambridge, MA.

FLETCHER CP (1973) The Capital Budgeting Decision. SM Thesis, MIT, Cambridge, MA.

FRIEDSAM LJ (1973) Top Executives' Viewpoints on Corporate Social Responsibilities. SM Thesis, MIT, Cambridge, MA.

GILMAN M (1973) An Investigation of Real Estate Investment Syndicates. SM Thesis, MIT, Cambridge, MA.

HAGENBUCH HW (1973) The Selection of Plant Locations Abroad. SM Thesis, MIT, Cambridge, MA.

HARTMAN RL (1973) Limits to Growth of Physics. SM Thesis, MIT, Cambridge, MA.

HELLSTROM DR (1973) Market Valuation and Price Movement. SM Thesis, MIT, Cambridge, MA.

HENEGEN ML (1973) Profit Centers at United Air Lines and its Eastern Division. SM Thesis, MIT, Cambridge, MA.

HOM GP (1973) Key Concepts in an Introductory Management Information Systems Course. SM Thesis, MIT, Cambridge, MA.

HURTADO-SANCHEZ L (1973) Short Interest and Monthly Excess Stock Returns. SM Thesis, MIT, Cambridge, MA.

KATZ MS (1973) Computer Networks as a Means for Satisfying an Organization's Computer Needs. SM Thesis, MIT, Cambridge, MA.

KRASNICK DB (1973) Simulation Based Analysis of the Courtroom. SM Thesis, MIT, Cambridge, MA.

KUSIATIN I (1973) The Strategic Behavior of American Defense Firms, 1946–1971—A Comparative Analysis. SM Thesis, MIT, Cambridge, MA.

LALL K (1973) An Introduction to the Modern Theory of Corporation Finance. SM Thesis, MIT, Cambridge, MA.

LINKAS T (1973) The Proposed Expansion of the Joslin Diabetes Foundation: Meeting Patient Needs. SM Thesis, MIT, Cambridge, MA.

MARKS HJ (1973) The Variable Rate Mortgage Instrument. SM Thesis, MIT, Cambridge, MA.

MOORE DB (1973) Health Care Delivery Within a Day Care System. SM Thesis, MIT, Cambridge, MA.

NIGG JJ (1973) Solid Waste Disposal System Management. SM Thesis, MIT, Cambridge, MA.

PETERSON VL (1973) Effects of Funding Policies on Technical Manpower Utilization. SM Thesis, MIT, Cambridge, MA.

POTCHEN EJ (1973) A Strategy to Study the Use of Radiology as an Information System in Patient Management. SM Thesis, MIT, Cambridge, MA.

RAJ R (1973) The Currency Risk Problem for a Multi National Enterprise. SM Thesis, MIT, Cambridge, MA.

ROUSSELIN A (1973) The Impact of International Diversification: A Study of the French Mutual Funds Industry. SM Thesis, MIT, Cambridge, MA.

RUDOLPH JD (1973) Prior Approval Programs in Medicaid: An Overview and Analysis. SM Thesis, MIT, Cambridge, MA.

SANCHEZ-BELL A (1973) Applications of a Dynamic Programming Model for the Stochastic Cash Balance Problem with Average Compensating Balance Requirements. SM Thesis. MIT, Cambridge, MA.

The following are obtainable at the Lockwood Library, State University of New York at Buffalo, 3435 Main Street, Buffalo, New York 14214, USA:

ACITO F (1976) Consumer Preferences for Health Care Services: An Exploratory Investigation. PhD Thesis, University of New York at Buffalo.

BERKOWITZ M (1976) Pricing and Production Decisions for Regulated Public Utilities. PhD Thesis, University of New York at Buffalo.

CASTELLANO J (1975) Influence of Worker Background and Work Values on Job Attitudes and Job Behavior. PhD Thesis, University of New York at Buffalo.

DEAL K (1975) A Differential Games Solution of the Problem of Determining the Optima Timing of Advertising Expenditures. PhD Thesis, University of New York at Buffalo.

GREENE F (1974) An Analysis of the Functioning of the Fact-Finding Process in Selected Schoolboard—Teacher Negotiation Impasse in New York State. PhD Thesis, University of New York at Buffalo.

KON S (1976) Overall Investment Performance: An Empirical Investigation of Timing and Selectivity. PhD Thesis, University of New York at Buffalo.

PALMER D (1975) Teachers, Researchers and Consultants: An Empirical Study of the Faculty in Collegiate Schools of Business. PhD Thesis, University of New York at Buffalo.

RAMSEY J (1976) The Research and Development to Operations Planning Process: A Strategic Analysis. PhD Thesis, University of New York at Buffalo.

SAVITT J (1976) Electric Power Usage and Regional Economic Development. PhD Thesis, University of New York at Buffalo.

SMITH L (1975) Dynamic Inventory Models for Perishable Commodities with a Fixed Life. PhD Thesis, University of New York at Buffalo.

VREDENBURGH D (1975) An Investigation of Relationships Between Structuring in Organizations and Individual Attitudes and Performance. PhD Thesis, University of New York at Buffalo.

YERKES R (1975) Variations in Adherence to Expressed Rules of Ethics by the accounting Profession and some Factors Associated with such Variations. PhD Thesis, University of New York at Buffalo.

Information for inclusion in this Section, presented in the appropriate form, should be sent to the Editor of Omega.

BLANK
PAGE

Vol. 5
1977

OMEGA

The International Journal of Management Science
Volume 5, Number 6, 1977

Contents

SAMUEL EILON	627	Editorial—More Against Optimization
JEAN-JACQUES SALOMON	635	The Contribution of the Social Sciences
RL ACKOFF	649	Towards Flexible Organizations: A Multidimensional Design
N MEADE	663	The Decline of Venture Capital
S EILON, GP COSMETATOS	673	A Profitability Model for Tactical Planning
HG JONES	689	The Management of Timber and Related Industries in Sweden
JDW MORECROFT, CDT WATSON-GANDY	699	Market Representation in the Motor Industry
W DAVIS, J TALAVAGE	709	Three-Level Models for Hierarchical Coordination
SP LADANY, DN BEDI	721	Dynamic Booking Rules for Flights with an Intermediate Stop
CC GALLAGHER	731	Production Technology and Management in Planned and Market Economies
LAL C JAGETIA	741	Memoranda
ROBERT M BROWN	746	
PG BENNETT	749	
	753	International Register of Theses and Dissertations

Special Notice

To avoid simultaneous, multiple submissions to journals having similar interests, the Editors of *Management Science*, the *Operational Research Quarterly*, *Operations Research*, *Omega* and *Naval Research Logistics Quarterly* have developed a system of exchange submission lists.

Because refereeing capability is a scarce resource, we cannot tolerate the duplication of effort that follows from multiple submissions. We hope that the elimination of multiple submissions will make the refereeing process more efficient, and thereby decrease the time required for refereeing. This is a major objective of all five publications. Other journals with similar interests may soon be included in this information interchange.

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 6, 1977. Pergamon Press. Printed in Great Britain

EDITORIAL

More Against Optimization

YOU ONLY have to open any journal in the field of OR (Operational Research) published during the last 20 years to realize how obsessed academic analysts have been (and, for that matter, still are) with optimization. OR is expected to advise management on the efficient use of limited resources, so the implied rationale goes, and what more efficient use can there be than the optimal use, namely the maximum utilization of the most valuable of the constrained available resources to ensure the maximum attainment of an objective set by management as its primary goal?

The fascination of finding an optimal solution to a given problem lies not only in the intellectual skill that may be involved, and not just in the obvious fact that the solution dominates all other alternatives in its attractiveness, but also in its ability to serve as a yardstick against which a given performance may be measured: instead of having to draw any conclusions about the performance of system *A* by comparing it with another system *B*, a comparison that requires some justification since the two systems are invariably subject to different operating conditions, it is possible to resort to a comparison of a given performance of system *A* with the best possible outcome if resources at the disposal of *A* are optimally utilized. In this way a true measure is obtained of the way in which the system is managed, relating it to an operating environment which is under the jurisdiction of the management in question, whereas a comparison with another system often involves variables outside management's control and therefore limits the scope for making judgements about managerial performance in either system. Thus, a measure of performance is devised which is both absolute and unique: it is absolute in that it provides an unambiguous and indisputable (for a given problem) comparison of what has been achieved with what can be achieved, and it is unique in that it is tailor-made for a particular system in a given environment.

The search for optimal solutions is not, therefore, merely an intellectual exercise pursued by purists for the fun of it; it is a reflection of the philosophy of economic man whose aspirations are to derive the maximum utility in a given set

of circumstances, by maximizing a measure of output for given inputs, or by minimizing the inputs required to attain a given output, or by optimizing some other criterion that relates outputs to inputs. Any managerial activity which is embedded in the concept of economic man must naturally turn to optimization as a means of determining desirable or future courses of action.

An interesting facet of the optimization process is that it is often the natural outcome of marginal analysis. In exploring possible ways by which given industrial or business operations can be improved, the analyst poses the obvious question: how would performance be affected if each of the products or activities were subject to a small incremental change of, say, one unit? He identifies the product or activity X for which an increase of one unit will yield the greatest positive benefit and naturally concludes that such an increase should be undertaken.

Now, if the basic data and relationships remain thereby unaffected, further incremental changes in X are called for, until a point is reached beyond which no further increase in output is possible owing to some constraining resource; the marginal analysis is then repeated, new products or activities are identified and the maximum allowable changes are advocated, and when eventually this process of iteration does not yield any further avenues for improvement, the end result represents the logical outcome of incrementalism. In linear systems this result is acknowledged as the global optimal solution; for non-linear relationships it may be a sub-optimum (or a 'local' optimum), but in either case it dominates both the current state of affairs and all the interim iterative solutions based on incremental changes. And where the analyst has reason to suspect that the solution is not globally optimal, he will proceed to search for it, for example by comparing sub-optima (each derived by an iterative analysis related to a different starting point).

There is another, and perhaps less overt, driving force behind the doctrine of optimization and that is the conviction among some circles that OR is in business to increase the domain of what are termed 'programmed decisions'. Such decisions are called for when problems occur with a relatively high degree of frequency and/or replication, so that developing a methodology for handling them is thought to be beneficial. The advantage of programmed decisions stems primarily from the fact that a manager does not have to spend time pondering on a routine problem and solve it afresh each time it emerges; not only does the developed methodology provide him with a means of ascertaining quickly what the answers to the recurring problems should be, but he can then delegate the routine work to lower echelons. Training personnel for the purpose can be greatly simplified to achieve a high level of performance and, furthermore, programmed decisions provide consistency, so that no matter who of the trained junior managers is faced with the problem of making the decision, the result is the same, if they all follow laid-down procedures. Consistency is a particularly important attribute in administrative systems in which the outcomes of decisions

Vol. 5
1977

affect people and where it is essential to ensure that individuals are seen to be treated in the same way, irrespective of who is assigned to handle their affairs, otherwise feelings of injustice and frustration are bound to ensue.

For a high degree of delegation to be effective in such circumstances, and for consistency in response to be maintained, a rule book needs to be developed. Hence, programmed decisions are none other than a set of prescriptions, which either provide clear instructions as to what actions should be taken in given situations or specify what method of analysis should be followed in order to arrive at decisions [1, Chap. 7 & 8]. In either case, the result is expected to be personality-independent and the prescriptions deprive the 'decision maker' of any true measure of choice.

Now, for such prescriptions to be effective, they have to be 'right', namely they have to remain valid in the face of possible challenge, and hence they should preferably be optimal. This is not to say that all procedures for programmed decisions in practice are indeed optimal, but clearly those which are based on optimization processes command greater confidence, from top management as well as from those who have to implement them, and the procedures are therefore likely to remain in force for a long time, unless the basic assumptions and relationships between variables are for some reason undermined.

Thus, we see the doctrine of optimization arising from three complimentary precepts: the first is associated with the philosophical aspirational principle of economic man, that it is the duty of the manager to do his best in economic terms and that he should both seek and implement optimal solutions; the second is the perceived inevitability of the pragmatic incrementalist, whose marginal analysis starts by concentrating on local explorations (the term 'local' is not used here in a geographical sense), from which he soon concludes that optimal solutions should be adopted; the third lies in the continuous encroachment of OR on the province of the decision-maker to help him delineate, classify and solve problems which he encounters in his work and thereby make more of his tasks prescriptive and programmable.

What should our attitude be to such notions? As for the stance of economic man and the general debate between the optimizer and the satisficer, I have commented elsewhere [2, Chap. 13] in some detail on the various issues involved, and I do not propose to repeat the arguments here. Suffice it to say that even when methods of solution are readily available to determine optima, I have some doubts about the advocacy of the optimizer (except within agreed narrow definitions of given problems). But how should one react in situations when such algorithms are not available? What happens then to the aspirations of economic man?

The first question that springs to mind is why does the optimal solution elude us? If the reason is that it is not entirely clear what we should optimize, that there are multi-criteria which we do not know how to reconcile or how to assign trade-offs to, that the input data are suspect or incomplete, in short that the

problem is in some sense ill-defined, then by definition there is no way in which the optimizer can be confident that a particular solution will satisfy the objectives of economic man, nor can he set up an absolute and unique yardstick by which that solution can be judged. And it was this reason that led Simon to consider the need for decision-makers to arrive at conclusions under the conditions of 'bounded rationality', where they "have not the wits to maximize"; he then propounded the philosophy of 'satisficing' as a means of describing how decision-makers actually behave in practice in seeking solutions that do not violate prescribed constraints [3], either because bounded rationality simply does not allow them to do any better or because it is in their interest to retain slack to allow themselves some room for manoeuvre in the future [1, p. 188]. In either case, satisficing gives rise to the use of heuristic methods, which aim to explore the problems in question and to arrive at good solutions. In the satisficing mode a 'good solution' is an acceptable solution, one which is not ostensibly dominated by others in terms of all the criteria employed, but which is difficult or impossible to place on an explicit single utility scale.

There may be another reason for not having algorithms that lead to optimal solutions, and that lies not in problems being ill-defined but in their mathematical complexity. Many combinatorial problems are of this type: their definition is clear cut, often dazzling in its simplicity, the input data crisp and deterministic, the objective function involving a single criterion, and yet search procedures are lengthy and cumbersome, often culminating in local optima and the global optimum impossible to ascertain without complete enumeration, or at best partial enumeration based on tree search methods guided by prodecures for establishing progressive bounds.

Vol. 5
1977

Here, too, heuristic methods are widely employed to determine 'good solutions' for a limited computational effort, but it is interesting to compare the reactions of many pro-optimization analysts to the development of heuristics in the two cases. The use of such methods for ill-defined problems is grudgingly accepted as something that has to be tolerated: the fault, so to speak, lies with those who choose to define the problem in the first place; it reflects their ignorance of what the problem really is, or their inability to specify what they want; no blame can possibly be attached to the analysts, nor can there be objective ways which allow the utility of any proposed solution (unless it is clearly dominated by another) to be properly evaluated. But for the second class of problems, such as the combinatorial problems mentioned above, the shortcomings do lie with the analysts; if an optimal solution to a given problem does exist, it is they who fail to discover it by efficient means, so that questions about the quality of heuristic solutions legitimately arise. As a result, I have discerned a growing hostility from the orthodox establishment (which is by and large optimization orientated) to heuristic methods, treating them as the poor man's tools of the trade, as expediencies that should be regarded with deep suspicion, because they consume effort that should be better spent on searching for analytical

algorithmic solutions. If there is an optimal solution to be found, the purist will not be content with second best.

I have little sympathy for these attitudes. Even in the case of ill-defined problems, where the single goal optimization is largely irrelevant, the fact that optimizers are prepared to tolerate satisficing is merely a reflection of their inability—at least temporarily—to do any better. But put them in charge of an ill-defined problem, or in a position to influence the choice of models or criteria, and they will inexorably find ways in which to enforce an optimization process, often with the mistaken view that it is better to optimize something than nothing. And as for the use of heuristics to find approximate solutions, I am all for it, particularly where it can be shown that the heuristics are efficient computationally and produce good solutions. I realize that such an attitude creates problems in making judgements about the efficacy and the relative merit of competing heuristics, and indeed that heuristics can become the lazy analyst's charter, but the relentless pursuit of ultimate global optima, a pursuit which often proves to involve rapidly diminishing returns at very high costs, must surely remain the province of a diminishing number (one hopes) of academics and should rank rather low in the priorities of those striving to solve real industrial problems.

And what about the incrementalist, the man who discerns an incremental benefit from an incremental change and then infers that a large benefit will accrue from a relatively large discrete change? The fallacy of optimization based on incrementalism lies in the assumption that where a change takes place, the relationships between the major variables in the system remain nonetheless unperturbed. The assumption may well hold when the changes in operating conditions or in product mix are marginal, but they rarely hold when comparatively large changes take place. Thus, optimization based on conventional incrementalism cannot be valid. The iterative process of marginal analysis must always be accompanied by an exercise for updating assumptions, constraints and relationships, with the result that an advocated course of action may well take the form of implementation in several stages, with benchmarks for adaptation and review in between, rather than a recommendation for a drastic discrete change in operating conditions.

There is clearly a dilemma here for the management scientist. If he is in business to advise and influence decision makers, if he is to act as a change-agent, if he is to make an impact on his environment within a reasonably short period of time, then he is bound to be concerned about the possibility of implementing big changes rather than small changes, since the latter are less likely to result in the desirable impact. And yet, for his analysis to be reliable, and for outcomes to comply with forecastable expectations within reasonable confidence limits, he must resort to marginal analysis and be content with relatively small incremental changes. Optimization based on incrementalism is certainly not the answer to his dilemma. We simply have to accept that among management scientists there are many who are concerned with tactics and whose

contribution to operational effectiveness may best lie along the path of incrementalism, both in analysis and in the implementation of change. Others may see themselves as strategists, concerned with major reorientation of a given system, and their case will stand on their assumptions as to how the system and its environment will react to drastic changes in operational conditions and mix of activities.

As for the relentless onslaught of programmed decisions, it has to be said that if consistency in performance is their prime purpose, then their internal logic and specific contents become secondary; the warm approbation that their development receives from top management is seen in the bureaucratization of numerous administrative processes and, in terms of its purpose, this process is perfectly understandable. But curious as it may first seem, the linking of programmed decisions to optimal solutions is less valid a proposition. It is based on the premise that if programmed decisions are rooted in a methodology of optimization they are bound to outperform decisions based on arbitrary evaluations or on intuition, and it proceeds to conclude that such programmable decisions should therefore be programmed.

I, for one, do not accept this premise. For a start, it totally ignores the possible reactions of decision-makers, whose discretion is thereby threatened, to the imposition of a prescribed way of life; the loss of freedom to use their own judgement and to weigh up alternatives means so much to some that they may even intentionally distort the statement of the problems that need to be solved and the measurement of the major parameters that govern them [1]. It is therefore questionable whether the optimization of programmed decisions genuinely relies on modelling reality, and the results may not necessarily be superior to those taken through the exercise of discretion. Furthermore, programmed decisions are, by their very nature, set against a static framework of circumstances for which responses are envisaged; they can rarely be made flexible enough to cater for the unexpected, and consequently they lack the important element of adaptive learning, which is a vital ingredient for the control of systems under dynamic conditions. Thus, while programmed decisions have a role to play, in relieving the manager of the tedium of routine, in ensuring consistency in response, and in training junior managers to man seemingly complicated control systems, their contribution is paramount under essentially static and predictable conditions. It must be recognized, however, that when such conditions are not met, the scope and effectiveness of enforced programmed decisions are bound to be limited. In particular, artificial exercises in attempting to structure ill-defined problems through the instrument of optimization can prove to be not only abortive but positively damaging.

All this is not to deny that in some special circumstances optimization has a role to play, but set against the vast canvas of managerial problems, particularly those that are ill-defined, that are in the strategic domain and that matter most, this role is quite out of proportion to the dominant position that optimization

Vol. 5
1977

has assumed in the OR literature. One can only conclude that this position is a reflection of the fact that the problems for which optimization is largely irrelevant are much more difficult to resolve and that OR has not yet accomplished enough in developing the methodology needed for the purpose.

SAMUEL EILON
Chief Editor

REFERENCES

1. EILON S (1971) *Management Control.* Macmillan, London.
2. EILON S (1977) *Aspects of Management.* Pergamon Press, Oxford.
3. SIMON MA (1961) *Administration Behavior.* Macmillan, New York.

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 6, 1977. Pergamon Press. Printed in Great Britain

The Contribution of the Social Sciences[1]

JEAN-JACQUES SALOMON[2]

Organisation for Economic Cooperation and Development, Paris

(*Received June 1977*)

The contribution which Social Science has made and may make to policy formulation in modern Europe is described and evaluated. The rapid economic growth of the post war period was accompanied by the growing impact of government economic policy but the contribution made by Social Science in most countries was limited. The questioning of growth and bureaucratization which has subsequently arisen offers new opportunities for Social Science in policy-making. At present, its contribution is handicapped by limitations of methodology, and in some circumstances by a lack of awareness of its value. A positive strategy is now required for the development and integration of Social Science into policy formulation.

IN THE last chapter of one of the most thorough studies ever written on Western European industrial development, *The Unbound Prometheus*, one finds this statement: "It is impossible in the present state of our knowledge to evaluate the parameters of economic development. Even the European experience, the one we have studied longest and presumably know best, is still in many areas *terra incognita*"[1]. Such a statement does not, of course, prevent the author from proffering some conclusions, but these are presented so carefully that he feels the need to apologise for his temerity.

If I refer to Landes's statement as a pretext to excuse my own temerity, it is not only because I will write about more recent developments in the European process of industrialisation than he did and therefore be exposed to a higher degree of risk in assessing and generalising facts and trends which are both evanescent and illustrative of specific situations; it is mainly because this statement underlines the very limits of the knowledge resulting from the social sciences.

[1] This paper was presented at a Symposium organized by the Science Policy Foundation in collaboration with The Commission of the European Communities in Brussels, 14–16 April 1977. Proceedings of this Symposium are due to be published by Pergamon Press.

[2] This paper represents only the views of the author and not necessarily those of the Organisation to which he belongs.

Indeed, this limitation will be part of my subject. Speaking of the 'contribution of social sciences' seems to imply that one can apply to social phenomena exactly the same techniques and logic as those which the natural sciences apply so successfully to matter and nature: in brief, that there is in the social field a technology equivalent to the engineering technology. I am perfectly aware of this school of thought which claims that, by virtue of quantitative methods, mathematical models and systems analysis, a 'social technology' may emerge as a result of the progress accomplished by research in the social sciences. Yet it is somewhat naive and positivist to expect that all social problems can be solved once and for all by a 'technological fix'. Moreover, so complex is the social process and so little coherent is our understanding of it, that even the best description, analysis and explanations are always only half of the answer and therefore ambiguous and partial tools for policy-making.

I must add another word of caution which has to deal with the very concept of 'social sciences'. Even in Europe, their designations, their division into academic disciplines and the links between them vary from university to university and from country to country. It was only in the 60's that they began to receive widespread recognition. However, the different classifications of the disciplines involved still correspond to different conceptions and research orientations. Even the more general terms remain ambiguous: social sciences, human sciences, behavioural sciences are understood differently in different countries and the boundaries between social sciences and the humanities or even philosophy are still imprecise. This is why I have to define what I mean here by social sciences. I mean the whole complex of disciplines which deal with the behaviour of man and social institutions, extending from economics and political science in sociology and psychology as well as geography, demography, anthropology, linguistics and even law and history insofar as they are concerned with social processes. This very broad conception of the social sciences implies all the more that I have to restrict my field: I will not speak of their contribution at large—which would be anyway an impossible task—but only insofar as they are associated with policy-making.

Not that the influence of social science can only be measured in terms of research undertaken in connection with policy-making. Especially in Western Europe, a great part, if not the greatest, of their influence, is felt outside this policy orientation and stems in particular, from sociologists–philosophers such as Michel Foucault in France or Jürgen Habermas in Germany, not to speak of the rival schools of thinking within the marxists, neo-marxists, maoists circles and the growing attention paid by the European 'intelligentsia' to the methodological models provided by linguistics (see Claude Levi-Strauss in anthropology) or by the generalisation of psycho-analytical concepts and their influence on the current debate everywhere on new types of education or new ways of taking care of people's health.

With all these words of caution, I can now begin.

Vol. 5
1977

I. THE GOLDEN ERA OF ECONOMIC PLANNING

It is necessary to distinguish two phases. If the concept of industrial society turns around the axis of economic growth, a change occurred precisely when economic growth came under question. The first phase extends from the end of the second world war to the mid-60's. This was the period of growth conceived almost as an aim in itself and during which some social sciences—first of all economics and demography, but also sociology and the so-called 'management sciences'—were called upon in policy-making as an instrument of planning. The rapid economic expansion during this period was exceptional by its length and its regularity. But even more striking was the fact that this expansion was the result not of the traditional 'free market forces', but in great part of the process of organised planning.

As Shonfield has shown, two factors have been predominant: the imperative of full employment and the acceleration of technological progress[5]. If most of the European nations have been able to meet the implications of these two factors, this economic success was not due to capitalism as such: indeed, the 'free market forces' have been associated with the growing influence of government on the economic system, to governmental policies for higher standards of living and social security and welfare which were accompanied by an effort to define a new environment that would make the best use of long-term investments.

This effort has generated new forms of organisation, new mechanisms of planning and prospective thinking in which the public and the private sectors were necessarily linked in some kind of coordinated, if not concerted action. "State intervention in the economy since the Great Depression," Shonfield says, "almost certainly represents the biggest single achievement of a non-military character in the history of modern government. The advance of social welfare and the maintenance of an orderly and stable economic structure during a period of exceptionally rapid growth have demonstrated the capacity of the state in new fields of economic and social management."[6]

True, this was a non-military achievement, but at the same time, if one thinks of the new structures that were created in most of the European government machineries in order to define and implement economic planning, one is tempted to say that these new structures kept something of the wartime organisation, namely the mobilisation of scientific advice at the highest level of the executive branch. The 'peaceful soldiers' were mainly composed of economists, demographers and few sociologists whose task was, as illustrated by the French Commissariat Général au Plan, to provide the forecasting models necessary to the preparation of decisions affecting economic development.

The horizon of social scientists involved in this policy-making process was limited both in terms of specialisation and influence. If demography, economics and especially econometrics, were predominant among the social disciplines which

the administrations decided to consult, it was simply because they were closest to the natural sciences pattern of applicability in view of their use of mathematical methods. In other words, in a time of quantitative economic preoccupations, the social sciences could not influence the policy-making process except through the legitimacy of quantitative demonstrations: economic and demographic indicators, measurement of Gross National Products, development of econometric models and of management techniques which in fact gave many people the idea—be they engaged in research or in politics—that social science is but a kind of social engineering, able to define and to nurture growth and to provide technical solutions to the problems that growth creates.

To continue the military image, the policy which was affected by social science research was as limited as if, in an army, the strategic objectives were to be defined only by the engineers. If it is true that social sciences do not lend themselves readily to any straightforward utilisation, nor that utilisation constitutes a criterion of utility, it remains that the influence of social sciences during this period was limited to the extent to which they were understood—and therefore to my mind often misunderstood—at best as the intellectual background and at worst as the technical tool for achieving quantitative economic development.

Vol. 5
1977

Such a broad generalisation needs, of course, to be qualified in view of the diversity of national experiences in Europe. In some countries such as France or Germany, if social sciences at large became increasingly 'fashionable', the activities and findings of disciplines other than the mathematics oriented ones were then still treated as irrelevant. It is fair to say that in these countries a great bulk of social science research remains far removed from the everyday business and concerns of the world of affairs—which does not mean that the way in which a society becomes aware of its problems is not affected, progressively, in the long term, by this kind of research. When one speaks of the 'utility' of the social sciences, one has to take into account, apart from those applications of knowledge which are really implemented, their more general, but also more diffuse influence on society. This indirect impact of the social sciences on individual and collective perceptions of constraints, opportunities and values may well be as good an illustration of their 'utility' as the examples where their results have been directly translated into policy-making.

However, in countries such as the Netherlands, Norway and Sweden, social sciences were early regarded as a necessary aspect of the whole process of the development of society. This view has permeated the build-up of social research as an instrument of local, regional and national development long before the crisis of the late 60's. In these countries, the social sciences were not only called in to help in the implementation of pre-defined economic policies, but also participated directly in the elaboration of political decisions themselves, which means that they were already agents of social planning and control of technical, economic and social development.

This is characteristic of the Netherlands, Norway and Sweden where there

are very few fields of public concern in which psychologists, sociologists, geographers or political scientists have not been invited to contribute their views, alongside the economist and the demographers. Having this example in mind of the smaller European countries, it may well be that the impact of social science is commensurable with the degree of consensus that is reached in a given society. And the smaller the country the higher seems to be this degree of consensus, and therefore easier the process by which the social sciences become centrally involved in policy-making and their findings integrated in the overall make-up of society.

Two examples of the early 60's will show this striking difference in both the utilisation and influence of social sciences. In Germany, studies of the workers' participation in management were launched under the sponsorships of interested industrialists and trade unions. "Both sides," remarks a subsequent report, "employers and unions, expected sociological research to provide arguments with which to support their own policy and demolish their adversary's". In the event, although much debate arose in academic and industrial circles concerning the significance and validity of the research findings, little impact was made on the central problem of the constitutional reform of the system of industrial ownership and control in the context of an industrial society dominated by traditional values and attitudes in management.

Conversely, the same problem of workers' participation in management has led in Norway to a social research programme whose findings have developed a considerable public debate which went far beyond the power relationship between management and labour: what was at stake was to find a solution not only to the problem of industrial democracy but to the related and not dissimilar problems of social integration and of political participation and action in the larger society[7].

These two different examples confirm how much the impact of social science research in its 'policy-relevant' aspects is linked with the appropriate interest and response on the part of not only social scientists and policy-makers, but also the public at large regarded as the ultimate user of both the research findings and their exploitation within policy-making.

Not that, once again, one should expect from social science, even if such a partnership is successfully established between researchers, policy-makers and the people, clear-cut answers that can be directly implemented. The impact of social science takes on many other paths, notably by throwing light on problems, by clarifying choices, by proposing alternative options or putting aside unattainable objectives. The resort to social science as a field of research can only be a means of eliciting, collecting and analysing information, of formulating and testing hypotheses, and of constructing plans for social and political action in the light of the results obtained. As it was so well put in the Examiners' OECD Report of French Social Science Policy, such research may point "the margin of uncertainty inherent, for a part, in the state of our knowledge, and for another

part, in the very nature of social problems (. . .) This is all we can ask of social science: it lights up the way ahead; it will not build it."[8]

But, if such a linkage with the ultimate users is missing or ignored, there is a real risk of a technocratic misuse of social science. We all know that industrial societies very often attempt to base their legitimacy on the resort to 'objective criteria for decision-making', and reference to science is frequently made in order either to postpone taking a decision, on the pretext that more research is necessary, or to give an added credibility to decisions which have been already taken.

In a sense, one could say that something of this kind happened in the late 60's when social science restricted to its quantitative aspects was invokedas an alibi to underestimate the social consequences of economic growth. The traditional mechanisms of economic planning were obviously ill-equipped for coping with these social consequences, since the mathematical tools as well as the predominant Keynesian theory could not take into account the significant changes that have occurred in individual and collective behaviour precisely as a result of the rapid rate of growth. Industrialisation and urbanisation had not only provided a material affluence; they have been accompanied by important inequalities and distortions between regions, between the city and the country, between socio-professional categories, not to speak of the differences between the rich and the poorer countries. Furthermore, these distortions were accentuated by the by-products of the race for economic growth, which indeed were not mastered, if one thinks only of the deterioration of the natural environment, or the growing difficulty for individuals and groups to adjust to the speed and related psychological, sociological and cultural stresses of technological and organisational change.

Vol. 5
1977

For instance, the obsolescence of skill and information in so many sectors of industry and organised social life not only created tensions between managers, engineers and workers or even in the relationship between generations, but also posed the basic problem of what kind of education is required to meet the new professional needs and how the continuation of learning through maturity can be institutionalised within the education system. In sum, after two decades of continuous growth, the quasi-military headquarters organised for economic planning was taken by surprise, as if, in advancing further and further ahead, in terms of quantitative achievements, they had forgotten the people 'behind the lines' whose qualitative aspirations were less and less satisfied.

II. THE SEARCH FOR QUALITATIVE GROWTH

This 'technical surprise' was the beginning of the second period which, in France started around 1968 in the stream of the students' uprising, and which crystallised as from 1970 in the whole debate on the questioning of growth. One

has often said that it is characteristic of industrial societies that they seek, through rationalisation, to anticipate, plan and organise their undertakings, to gauge the risks and predict and control the repercussions. But either the tools which were supposed to be provided by the social sciences were lacking or only partially utilised, or the problems and the stake revealed themselves as being more complex than the technocrats believed.

Whatever the answer is to this failure of industrial societies, it remains that this second period emerged with the growing awareness that growth has in the last analysis a social objective and requires a system of values and criteria on the basis of which economic considerations and decisions can be weighed in the light of the human and social factors involved.

On one side, economic growth as such appears more and more inadequate to meet the growing qualitative aspirations. Redistribution of the benefits and disadvantages of progress become therefore one of the major problems which industrial societies—in Europe as well as elsewhere—will have to overcome. It is in this framework that most European countries have started to develop a new system of socio-demographic statistics aimed at a better understanding of the conditions for welfare which are not produced via traditional market forces; furthermore, the new awareness of the limitations of GNP as an adequate measure of 'output' has led to attempts towards a definition of 'social indicators' that could provide decision-makers with more balanced tools for economic planning.

At the same time, there is a growing amount of social science research being conducted by government agencies as well as universities on problems such as urban development, environmental pollution, health services, transportation systems, migratory workers, the reform of public administration, whose common feature is that the weight of the economists as such in the research teams and institutions is diminishing and the need for interdisciplinary research more and more recognised through the building-up of new types of research institutions favouring exchanges and cross-fertilisation between diverse disciplines, ranging from the social to the natural sciences.

On the other hand, one should not over-estimate this trend towards a greater resort to the social sciences in general, for two reasons: first, there is more than ever the risk of expecting from them more than they can actually offer in terms of scientific analysis and understanding of the various factors involved; second, time is short in view of the accumulated difficulties that industrial societies have to face. This sense of living in an interregnum and even in a crisis, is sufficiently well grounded to be more than intuitively based: the question posed as to the limits to growth, the multiplicity of models of saturation, if not the prophecies of doom and finally the 1973 energy crisis accompanied both by an accelerated rate of inflation and rising unemployment show that the process of development which was considered only a decade ago as a quasi-mechanistic process, can no longer be taken for granted and that industrial societies, in spite of their

scientific and technological capabilities, are no less vulnerable than human beings.

Many of the symptoms of these difficulties were already set out in the so-called Brooks report, *Science, Growth and Society*, published by OECD in 1971. Because of the impact it had, especially in calling attention to the need to promote social science research and to ensure "a much closer relationship between policies for science and technology and all socio-economic concerns and governmental responsibilities than has existed in the past", it is worth quoting some of its conclusions. For instance, the report stressed—1 year before the Meadows' Report, on *The Limits to Growth!*—that many aspects of developed societies are approaching a condition that may be described as the precursor of saturation, in the sense that things cannot go on growing much longer in some lines without reaching fairly fundamental limits.

"Indications of saturation are present in total population, pollution of the environment, in the size of urban conglomerations, in traffic, in information overload impinging on the individual, even in higher education and, perhaps, in the view of some people, the production of new knowledge. This does not mean that growth will stop in the next decade, or even that the time limits for growth in the future can now be foreseen in every instance, but only that a declining rate of growth is foreseeable within the lifetime of people now alive. In a society now accustomed to growth in almost all its aspects during the last 300 years, this is something quite new, something that will require considerable adjustment"[2]. And it is true that already since 1974 industrial societies have to learn to live with a growth rate which levels off rather than increases.

The report went on to say that in practice the changes brought about by economic growth are no longer spread out over long enough periods to deaden the impact of change. Whatever measures may have been taken to attenuate the harmful consequences of the 'by-products' of growth, they have not been enough to counterbalance the social costs which may, in the long run, appear to be too high a price to pay for the advantages derived. "In the face of the results achieved by 25 years of uninterrupted economic growth, and of the technical progress it has permitted, as exemplified by the voyage to the moon, the man in the street naturally tends to think that vigorous effort and judicious allocation of resources are enough to put us on the right lines towards an all-sufficient solution. This does not seem to be true: the destruction of the environment, the disequilibrium of the ecosystem, are not fortuitous events; they are linked with the whole course of our civilisation."[3]

"A crisis of civilisation": the very success of their technical achievements induces industrial societies to examine their future, to call in question the goals they set themselves and to ponder upon the limitations of the economic system in which they are developing. It is not enough to stake everything on growth, it is also essential to master the course of that growth. But, as the Brooks Report emphasises, the instruments available to economic science for that purpose are

Vol. 5
1977

insufficient, inadequate and uncertain. "Not enough is known about the complete technico-economic processes that are transforming society to give simple answers to the many problems they raise. The science of economics, despite all the refinements it has undergone in this century, has not been able to give policy-makers the kinds of advice they need. A more comprehensive approach is required, and it should be one of the tasks of future science policy to encourage through study of the interactions between what we traditionally called economic factors and the structural changes due to the extraordinary advances of science and technology in current history"[4].

This is why social aspects must be fully taken into account in the study and implementation of national as well as specific sectoral policies. Science policy in particular, but also overall governmental policy, should therefore be more concerned than in the past with research effort in the social sciences and pay greater attention to the application of these disciplines and their association with the natural sciences and engineering. From this approach, the function of social science in industrial societies is seen as a recourse for the many problems which growth and the uncontrolled exploitation of technology have brought about or have at least made more acute.

III. THE CHALLENGE OF IMPLEMENTATION

The critical problem facing our industrial societies is how to translate into action this new awareness. Obviously this can only be done at the concrete level of structures, processes and policies, that of specific sectoral concerns such as welfare, corporate, industrial or environment policies, community services, problems such as health, transportation, communication and urban amenity. In dealing with the problems at this level, it is only through the process of fact finding, research and comparisons between national experiences that we can hope to create a better linkage between what is, in effect, the machinery of society and the various wheels making it turn, between the overall guidance given by government and the way in which the social process operates, the various institutions which run it, government, industry, university, etc.

In the European context, this emphasis on sectoral concerns takes place at the very moment when the prestige, as well as the effectiveness, of governmental planning have declined. Indeed, the process of planning has tended to be no more than a prospective thinking exercise, whereas the need for a better understanding of the socio-economic factors which condition the evolution of industrial societies is more pressing than ever. The questioning of the concept of growth conceived of in its quantitative aspects has in fact created a vacuum rather than contributed to the extension of the scope of economic planning. For most European countries, the abrupt change in the economic background, accelerated if not provoked, by the energy crisis, has found traditional planning

mechanisms so little prepared that the great bulk of policy-making is now concerned with catching up with short-term constraints rather than with anticipating and organising the future. A well-known European statesman has said that he now can but "pilot from day-to-day the unforeseeable" and this formula is an echo of the former trend in the 60's towards forecasting and long-term planning. The fact is that, if everybody in Europe has a sense that an era is ending, there is little agreement as to what may be ahead.

And this is the paradoxical situation; long-term decisions need to be taken, but the tools that should become practical instruments for policy-making are lacking or suffering from methodological limitations. In supporting economic statistics, governments not only were able to be better equipped for the understanding and day-to-day management of economic policies, they were also in a position to influence the progress of economics as a scientific discipline.

The same interest with respect to social science information gathering should develop on the European scale and through concerted efforts as a guarantee of both the scientific progress of the disciplines and the opportunities for their application. In this regard, one must admit that if the social sciences are 'theory rich' they still are statistically weak.

Take, for example, the case of technology assessment or environmental impact statements, two fields where policy-making needs to rely more and more upon the findings of social science research. There is no objective way of measurement in part because there is no social statistical basis at all and therefore no common ground for a consensus as to the decisions that must be taken in either field. In the case of technology assessment, will we learn to measure unknown changes against unknown values? Yet, we are still unable to assess the social consequences of past technologies against our current conflicting values! In the case of environmental impact statements, nobody knows how to implement them in order to balance a consensus on the means against broader social preferences.

From this and so many other points of view, social science will most likely have more and more impact on the shaping of the future of industrial society in Europe. There are at least two reasons for supporting this view: the technical 'intelligentsia' is no longer confined to the scientist, the engineer and the manager, it includes also the growing number of professional social scientists whose influence is felt not only in bureaucratic and administrative circles, but also—and more and more—in the electoral system and the political process. At the same time, these new 'technicians' represent indispensable administrative advice and sometimes staff for politicians, union leaders or consumers' groups.

The clamour against bureaucracy and the desire for participation and more public control of the decisions affecting people's lives and fates will certainly be mediated through social science research activities. The short-comings and pitfalls of economic planning in the 60's have shown the need for more discriminating and coherent courses of collective action. Viewed from this aspect the social

Vol. 5
1977

science impact on industrial society will be to provide better public enlightenment and scientific material for decision-makers concerning the problems involved in the aims they pursue.

In most European countries the problem of the proper working relationship between social science research workers and institutions and the apparatus of government has already led to various institutional solutions. It is against this background that the OECD Ministers of Science had asked the Committee for Scientific and Technological Policy and the Secretariat, as a follow-up to the recommendations of the Brooks Report, to launch a programme aimed at understanding the problems involved in the closer association of the social sciences with governmental policy-making. This trend in most countries towards the definition and implementation of social science policies—namely policies both for a better development of the various disciplines involved and a more coherent utilisation of their results—will offer a natural conclusion to what I have tried to express in this paper.

In the 60's science policies were mainly technology-oriented in the stream of a conception of growth that was narrowly implemented in terms of quantitative targets. The challenge before the 70's expressed in the growing awareness of the complexity of the problems facing industrial societies calls urgently for science policies which will not be restricted to natural science and technology, but aimed more and more at probing the efforts of various scientific disciplines. The very fact that industrial societies seek to organise their development in a more rational manner, demands and will do so increasingly, much greater utilisation of the social sciences. At the same time, the fact that the analytical tools at their disposal are still insufficient to guarantee the efficacy of policies, indicates clearly how much remains to be done before knowledge in this field will provide an adequate guide to the measurement, anticipation and planning of social change.

Therefore the expectations raised by a closer association of social sciences to policy-making are linked in great part with the internal progress of the disciplines, with the amount of qualified, professional manpower and the amount of financial resources at the disposal, with their adaptation of the structures of research institutions—and last, but not least, with the status accorded to social sciences and the extent to which policy-makers are willing to pay attention to their research findings. Industrial societies are knowledge-based and their ability to innovate and adjust depend in great part on the information, analysis and critical evaluation provided by the social sciences, in relation to policy-making as well as in the framework of scholarly work detached from any policy relevance. However, the extent to which social sciences can be applied is also a function of the will of those in authority to support the fact finding and research process and to face the challenge of knowledge.

I remember that when OECD conducted an examination of science policy in the United States, we had the privilege of attending a meeting of the President's

Science Advisory Committee, and at this meeting one of our examiners, Mr. Théo Lefevre, former Prime Minister of Belgium, observed that in mainland Europe there was no radical distinction between the natural and social sciences: the German concept of *Wissenschaft*, the French notion of *science* or even the Russian notion of *Nauka*, covered all the disciplines of scholarly research and investigation. Why, Mr. Théo Lefevre asked, was there no representative of the social sciences on the President's Science Advisory Committee? At this one of the committee members, a famous physicist and Nobel laureate, exclaimed: "If ever there is a social scientist on this committee, I shall quit!"

That was in 1966. Two years later there was a representative of the social sciences on that committee, and one year later there was yet another. Today the National Science Foundation has a deputy director for the social sciences; and it is likely that the new President's Science Advisory Committee, restored under Mr Carter, will have several representatives of the social sciences on it. For once, the United States, which was so often looking far ahead of Europe in science policy, drew inspiration from European experience and not the other way around.

Indeed, the policies for the development and utilisation of the social sciences are often too synonymous or confused with social policy. But a social policy is only one aspect of the problems raised by a policy for and through the social sciences. Insofar as science policy is concerned with strengthening the contribution of scientific knowledge to social and economic development, there are three basic needs: to ensure, first, that the social dimension is taken into account in technological policies; second, that the social research potential is sufficiently responsive to tackle specific problems which preoccupy society; and third, that the social science community and its institutions enjoy the best conditions for improving the development of their multivarious research fronts. In brief, we have reached a point where long-term policies must be defined and implemented —i.e. a strategy, and no longer the partial measures which have hitherto characterised social science research.

Vol. 5
1977

This does not mean of course, that relations between social scientists and decision makers will not be coloured by cross-purpose dialogue and misunderstanding. Decision makers expect technical solutions, or recipes, for their problems; but for these kind of problems neither recipes nor technical solutions exist. It is true that, more often than not, social scientists contrive to complicate the issues to the point of obscuring the object of their investigations. Take the classic example of education: if the growth of the student population has not meant, in most of our countries, the democratisation which was hoped for, research findings over the last 20 years have added to the difficulty of assessing such a consequence. The problem has seemed to grow more and more complex, and we know only that education reform will have little effect in promoting equality of opportunity so long as the social structures remain unaltered. One could take examples from other fields, but none shows better than education

how the application of research findings in this field is never politically neutral, and how social science research tends to affect power relationships between the social partners rather than furnishing ready-made solutions.

I do not say that the social sciences have more claims to define ends for society than have any other disciplines; research findings are not to be confused with decisions. But they can help to define the conditions under which one defines an attitude towards society's goals and in this sense they help to clarify the alternatives before us. To conduct a society without struggle, passion and crisis is an end which is beyond the scope of any science. But if we know that politics, more often than not, is stronger than rationality and sometimes does itself upset rationality, we also know that there is no substitute for rational knowledge and action in the management of modern societies.

REFERENCES

1. LANDES DS (1969) *The Unbound Prometheus—Technological Change and Industrial Development in Western Europe from* 1750 *to the Present*, p. 539. Cambridge University Press, London.
2. *Science, Growth and Society—A New Perspective* (1971) Report of the Secretary General's *Ad Hoc* Group on New Concepts of Science Policy, OECD, p. 21.
3. *Science, Growth and Society—A New Perspective* (1971) Report of the Secretary General's *Ad Hoc* Group on New Concepts of Science Policy, OECD, p. 31.
4. *Science, Growth and Society—A New Perspective* (1971) Report of the Secretary General's *Ad Hoc* Group on New Concepts of Science Policy, OECD pp. 33–34.
5. SHONFIELD A (1965) *Modern Capitalism—The Changing Balance of Public and Private Power*. Oxford Press.
6. SHONFIELD A (1969) Business in the twenty-first century, in *Daedalus*, p. 92. Perspectives on Business, Cambridge, Ma.
7. For these examples, see *The Social Sciences and the Policies of Governments* (1966) OECD, pp. 55–57.
8. Social Sciences Policy, France. Examiner's Report (1975). OECD, Paris, pp. 229–230.

ADDRESS FOR CORRESPONDENCE: *Jean-Jacques Salomon Esq., Chef de la Division des Politiques de la Science, OCDE, Direction de la Science, de la Technologie et de l'Industrie, 2, Rue André-Pascal, 75775 Paris Cedex 16, France.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No.6, 1977. Pergamon Press. Printed in Great Britain

Towards Flexible Organizations: A Multidimensional Design

RL ACKOFF

The University of Pennsylvania

Received May 1977

The author considers one way by which the readiness, willingness, and ability of organizations to change themselves can be increased. Specifically, he considers how organizations can be designed to be more *flexible* and, therefore, more capable of being changed and changing themselves. Flexibility does not guarantee adaptability but it is essential for it.

IT IS apparent that the rates of technological and social change are increasing. Alvin Toffler [1], among others, has argued that our inability to cope with such acceleration, an inability he calls 'Future Shock,' is the principal producer of the many social, institutional, and organizational crises that confront us. Although those who point out the crises do so in harmony, there is considerable discord among them when they come to suggesting cures. It is not surprising, therefore, that enterprises, institutions, and societies continue to be managed and organized much as they were before awareness of their bombardment by change was widespread.

Most institutions and enterprises seek what Donald Schon [2] has called a 'stable state'. Their resistance to change tends to be proportional to the need for it. The more turbulent is their environment, the more stable is the equilibrium they seek. They fail to realize that the only equilibrium that can be obtained in a turbulent environment—like that obtained by a ship in a rough sea—is dynamic. A turbulent environment requires that institutions and enterprises be ready, willing, and able to change themselves. Without changing themselves they cannot adapt effectively to external change.

The design of organizational structure is part of what some call 'social architecture'. In architectural terms what is presented here is a sketch rather than a set of working drawings which can be used to build an actual structure. It is a presentation of an idea on which many different working drawings can be based, each meeting the unique as well as the shared requirements of different organizations in different states.

INPUTS AND OUTPUTS

Organizations are purposeful systems some of whose parts are purposeful individuals among whom there is a functional division of labor. Their purposefulness entails choice of *ends* (desired outcomes: objectives and goals) and *means* (courses of action). Courses of action necessarily involve the use of resources (inputs) to produce goods or services (outputs) that are intended to be of greater value to those who consume them than are the inputs. The resources consumed include manpower, materials, and energy, plant and equipment, and/or money. This is as true of organizations that are not-for-profit as for those that are, as true for those that are publicly owned as for those that are privately owned.

Traditionally, organizational structure is taken to deal with two types of relationship: (1) *responsibility*, who is responsible for what, and (2) *authority*, who reports to whom. Structure so conceived is conventionally represented by a two-dimensional tree in which boxes represent responsibilities and altitude and lines represent respectively the loci and flow of authority.

Our concept of organizational structure is greatly constrained by this conventional way of representing it. There is no necessity for this unconsciously self-imposed constraint or even the more general one: that organizational structure be represented graphically. Because the organizational structure presented here cannot be represented adequately in two dimensions, I call it 'multidimensional' or 'MD'.

The conventional representation of an organization's structure does not show either what inputs flow into what outputs or what means are used for what ends. An input–output matrix is seldom used to describe an organization's structure; a means–ends matrix virtually never is. Using them, however, can reveal new and more flexible ways of structuring organizations. In developing this point I use for illustration a typical privately owned corporation engaged in producing goods. Other types of organization are used in subsequent illustrations.

An organization's outputs can be used to define its ends; for example, outputs classified by product type or brands. I will refer to an organizational unit whose responsibility it is to provide a product or service to consumers outside the organization as an *output activity* or *program*. $P_1, P_2, \ldots, P_K$ will be used to represent these.

Activities are the means employed by programs. They can usually be divided into *operations*, ones that directly affect the nature or availability of the company's output; and *services*, ones that are required by programs or operations. Typical operations ($O_1, O_2, \ldots, O_m$) are raw-material purchasing, transportation, production, distribution, and marketing. Typical services ($S_1, S_3, \ldots, S_n$) are provided by accounting, data processing, maintenance, industrial relations, financial, personnel, and legal units.

A precise and completely general distinction between operations and services

is difficult to make, but doing so is not essential to the concept of organizational structure developed here.

The ways activities feed into programs and into activities themselves can be displayed in such tables as are shown in Figs. 1a and 1b. The output of each activity may be consumed by programs and other activities; additionally by the executive function (discussed below), by external consumers (also discussed below), and by that activity itself (for example, an accounting unit may prepare its own accounts).

Programs can be divided into subprograms which are defined, for example, by type of customer (industrial or individual), by geographic region supplied or serviced, by brands, and so on. Subprograms can also be subdivided.

Activities can be similarly divided. For example, a manufacturing operation may be divided into parts production, subassembly, and assembly; and each of these may be further subdivided. The same is true for service units. For example, a financial department can be divided into payroll, accounts receivable, accounts payable, and records sections.

If the number of programs plus line and support activities exceed the number an executive can reasonably coordinate, coordinators may be required within the executive function (See Fig. 2.). More than one coordinator or coordinating unit may be required on any one dimension. If the number of coordinators justify it, supra-coordinators or coordinating units may be required. It should be emphasized that in this context 'coordination' means *coordination*, not *direction*.

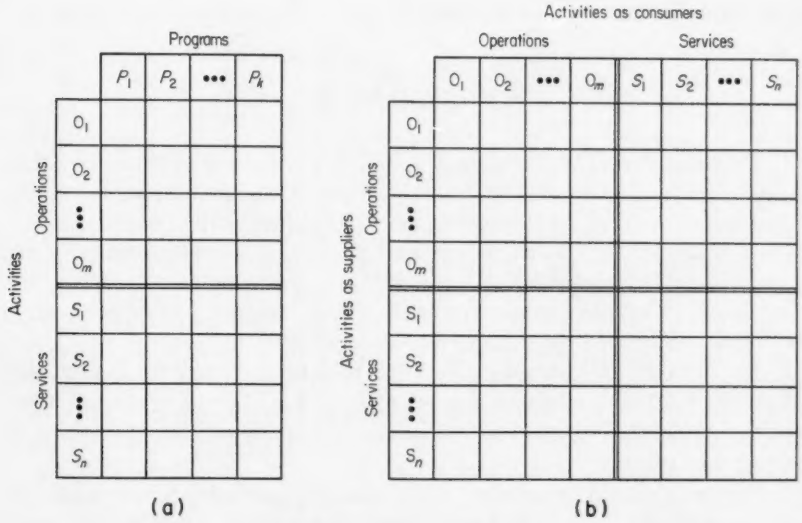


FIG. 1. *Activities and programs.*

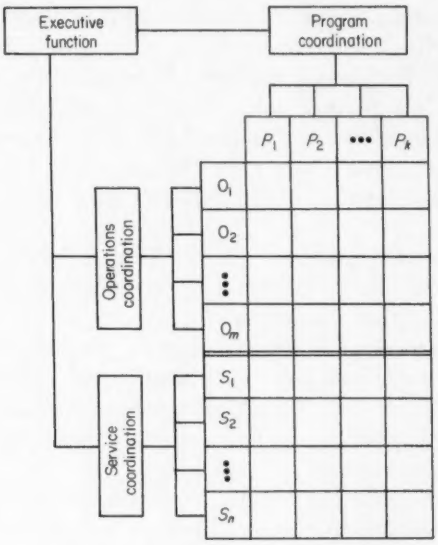


FIG. 2. *Coordination*

Vol. 5
1977

It may well be provided by a group made up of the heads of the coordinated units and the executive.

Now consider how an MD organization can be operated so as to maximize its flexibility and responsiveness to changing internal and external conditions.

PROGRAMS

1. Programs should be formulated so that they require more than one source of inputs. These sources may be internal or external. A program which requires only one such source should be combined with another so this condition is met. As will be seen, the more activities a program requires, the more flexibly it can be managed. The corresponding requirement should be imposed on operations. They should be designed so as to be required by more than one program and, in the case of services, by more than one line activity.

2. The objective of each program should be defined so that the benefits that can be derived from it can be measured. Further, because its costs can be measured, a specified function of its benefits and costs can and should be used to measure its performance.

3. Programs should be provided with managers and their personal staffs, *but no other personnel.* They should not be organizational units in the usual sense. They should be managements whose function is to specify, acquire, coordinae,

and control such input activities of others as are required to attain their objectives. They should have no authority over those who supply these services but they should be able to control them in a way revealed in paragraph 6 below.

4. Programs should be provided with *no plant or equipment*; they require no investment. Therefore, 'return on investment' is not an appropriate measure of their performance.

5. Programs require and should be provided with operating capital. They should receive at least as much of the income they generate as is required to cover their operating cost. If they do not generate sufficient capital for this purpose they can either borrow money to be paid back subsequently or be subsidized. They should be subsidized only when their actual or expected benefits exceed their costs and their costs exceed their income.

6. Programs can buy goods and services from either internal activity units or external suppliers, *as they see fit*. It is through their purchasing power that they control and coordinate the activities of their suppliers.

For practical reasons it may be necessary to constrain this freedom of choice; for example, because proprietary information is involved. The formulation or technology of a product, or its production process, may be 'classified' information hence the company may be unwilling to have it produced externally. Wherever such a constraint is imposed on a program by the executive function the saving, if any, that could be obtained by breaking the constraint should be determined. This should be paid to the program by the excutive function. If the constraint is self-imposed by the program, there would be no such payment. It is possible, of course, that there is no suitable external source, leaving the program no choice.

These conditions assure ease of evaluating and comparing program performance and thus make it possible to allocate resources to them in a way that reflects both their productivity and their relative importance to the organization as a whole. This is as true in not-for-profit and public organizations as it is in for-profit and private organizations.

Programs are ends-oriented units. Their managers should be concerned exclusively with attainment of the objectives that define them, not with providing the means required for their pursuit. Unless constrained by the executive function, they can buy these either from within or without, wherever they can obtain the best buy. They pay for services obtained from within just as they do for those obtained from without. This means that internal suppliers must compete with external supplies for the 'business' of the programs.

Because programs require no investment and only a few people, they can easily be added, subtracted, or otherwise modified. This gives the organization great flexibility.

The performance measures applied to subprograms can and should be constructed so that their sum is equal to the measure of performance applied to the programs of which they are part. Program directors can allocate their

resources to subprograms just as the executive or program coordinator does to programs.

Programs are what might be called the 'leading edge' of MD organizations. They are what organizations are about. All of their other activities are justified or not primarily in terms of their contribution to programs.

ACTIVITY UNITS

1. Activity units are provided with personnel, plant, and equipment as required. These are treated as investment which are expected to provide the organization with an adequate return.

2. They can sell their goods and/or services to programs and/or to external customers *as they see fit* unless explicitly constrained from, and compensated for not doing so by the executive function. Therefore, they operate much as independent enterprises do whether their parent organization is publicly or privately owned.

3. These units receive no subsidies, only investment. They are expected to generate their own operating capital from sale of their output. They can borrow money internally but are expected to pay back such loans. If they are not self-supporting, the parent organization may either allow them to go out of business, sell them, or take them over and reorganize them.

Vol. 5
1977

As previously indicated, if the parent organization wants an activity unit to withhold its products or services from external customers, it should pay the unit for doing so. The amount involved should be at least as great as the profits that the unit would have made from the curtailed sales of its output. This makes it possible to account for means and ends separately and to focus each director's attention on a well-defined responsibility. It also makes it possible to put financial responsibility for executive decisions on the executive function and to account meaningfully for this function. The executive function can receive a return on its investment and loans, share unit profits, and incur cost.

Activity units and programs turn over a portion of their profits to the parent organization but they should be allowed to retain some for self-determined reinvestment or other appropriate self-development activities.

Programs and activity units are virtually autonomous. As long as they perform satisfactorily, there is no need for intervention from above. The same is true for subprograms and subactivities.

If the executive observes that a service or product not provided internally is consumed heavily within the organization he may add an appropriate activity unit or extend an existing one. Any existing activity unit may do the same on its own initiative if it can do so without subsidy.

If an activity unit is used heavily by external but not internal consumers, the executive should want to know why. Either the programs are not purchasing wisely or they no longer have need for the goods or services provided by the

activity unit. In the latter case the activity unit may either be converted into a program, be sold, or be discontinued. On the other hand, if an activity unit has little or no external sales, and it has a capacity to meet an external demand that can be shown to exist, the executive should also want to know why. Chances are that it is not being managed competitively.

If an activity unit is used heavily by external consumers the executive function may well decide to create a program defined by the unit's output. For example, if an internal data-processing unit develops a large external business, the executive function may decide to put the organization as a whole into this business by creating a data-processing program. Unless constrained from doing so by the executive function, the data-processing activity unit would continue to service external customers and the program could buy similar services from external suppliers.

Activity units can also have an MD organizational design. For example, a Research and Development Unit, as well as an independent company of this type can set up its projects as programs. Its operating units could consist of disciplinary groups such as mathematicians, physicists, chemists, and so on. Its services could consist of drafting, computation, editing, and so on. Because any part of an organization can have an MD design the whole need not be so designed. This makes it possible to convert a traditional organization to the MD type in stages, starting at the bottom, the top, or somewhere in between.

Activity units can sell their services to other activity units as well as to programs, and to the executive function as well. For example, an accounting or personnel unit could be used by every other unit of the organization. On the other hand, an accounting or personnel unit may use the marketing unit to sell its services externally.

As previously noted, programs and activity units can be subdivided using product classes, types of consumer, geographical regions, and so on. Where consumers of program outputs are numerous and widely dispersed, geography can be used in another and unusual way: as an additional organizational dimension. (See Fig. 3.) Regional representatives (R_1, R_2, . . . , R_n) should have no managerial responsibility for either programs or activities. *They should serve as representatives (advocates) of those who are affected by the outputs and activities of the organization as a whole.* They are *advocates* serving external stakeholders. They can evaluate programs and activities in each region from the point-of-view of those *outside* the organization who are affected by them. These evaluations can be fed back to the organization's executive, coordinators, or unit directors.

By calling together the regional representatives, any program director can obtain an overall picture of his program's performance across the area it serves, and in each region. This enables him to allocate his resources across regions more effectively.

Geography is not the only criterion that can be used to classify external

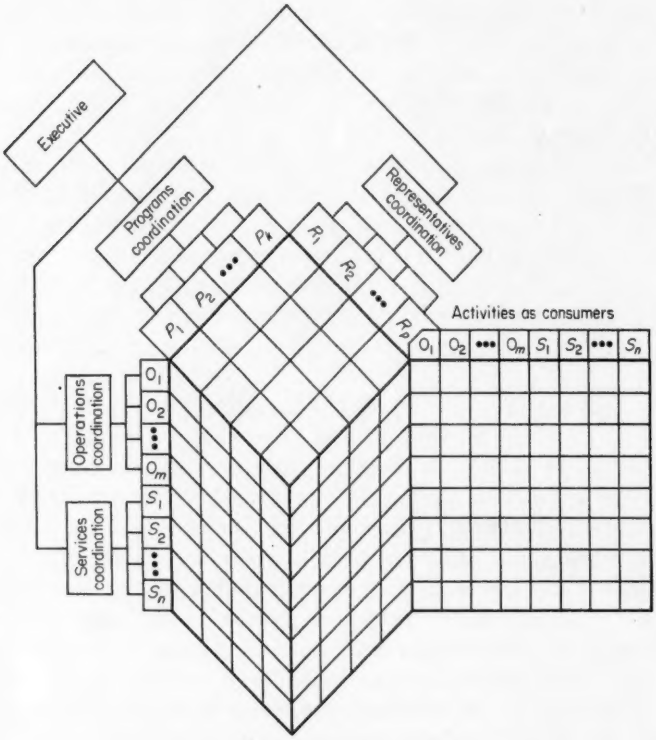


FIG. 3. *A multi-dimensional design*

stakeholders. Many other criteria can be used. For example, a company that supplies a variety of different industries—say, with lubricants—may find it useful to have advocates for industries rather than regions. They might be designated as automotive, aerospace, machine tool, and so on. A public-service organization may define its advocates' responsibility by the socio-economic characteristics of its customers.

SOME VARIATIONS

The MD organizational design presented above is a theme around which many variations can be composed, variations that are responsive to the unique conditions that pertain in each case. I mention only two here.

A recently prepared design of the national health and welfare services in Iran involves a 3D (programs, operations, and services) organizational structure for each of the regions into which the country is divided. The smaller regions are

Vol. 5
1977

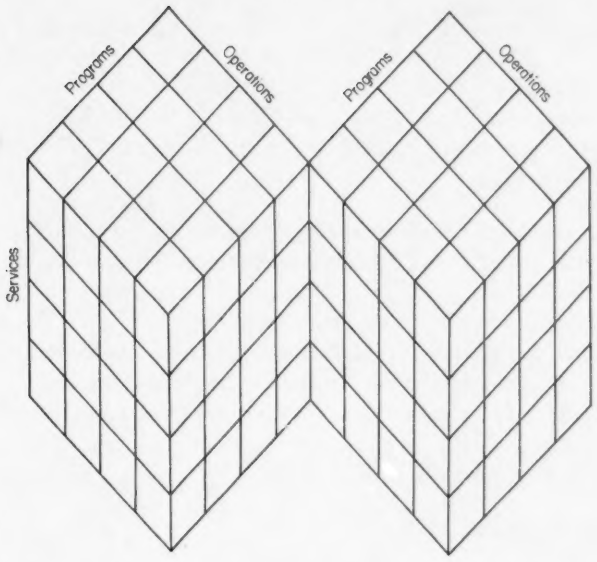


FIG. 4. *Two connected parts*

not provided with all the operations and services they require but these can be bought from units of larger neighboring regions. The regional executives report directly to the minister. This design, orginally stimulated by a desire to decentralize management and to be more responsive to varying local needs, also provides the flexibility required in a rapidly changing society.

Another variant is to be found in the organization of the Industrial Management Institute in Tehran which designed the health and welfare system described above. The Institute is divided into two parts, each a 2D matrix of programs and operations. They share common services. (See Fig. 4.) Operations can also be used as a common third dimension to any number—not necessarily two—of 2D matrices.

SOME EXAMPLES

MD organizational design is not restricted to privately owned enterprises; it is applicable to any type of organization and is, if anything, easier to apply to public than to private organizations. To show this I use two examples of very different types of organization.

A government agency

The first example is based on a study done of and for CONASUPO, the

National Basic Commodity Agency of the Mexican government. This agency operates with a budget slightly under $3 billion per year, employs about 14,000 persons, and is responsible for the production, distribution, processing and sale of subsistence goods, particularly food.

The five programs incorporated into the design were defined by the organization's five objectives:

(1) to increase income of low-income producers of subsistence goods,
(2) to increase production of subsistence goods and the productivity of such production efforts,
(3) to improve the quality of subsistence goods,
(4) to increase availability of (a) subsistence goods to low-income consumers, and (b) supplies and services required to produce them, and
(5) to increase the purchasing power of (a) low-income consumers for subsistence goods, and (b) low-income producers of such goods for the supplies and services required to produce them.

Seven operations were identified:

(1) buying (raw materials);
(2) transporting;
(3) producing;
(4) storing;
(5) selling, leasing, and lending;
(6) financing and insuring, and
(7) educating and developing intended beneficiaries.

Vol. 5
1977

Nine services were identified:

(1) information services;
(2) legal services;
(3) accounting and control;
(4) research and development;
(5) personnel services;
(6) purchasing (supplies and services), personnel transportation, and maintenance;
(7) public relations;
(8) construction and production of facilities and equipment, and
(9) general services.

Programs were further broken down using *subsistence classes* (food, clothing, health products, shelter and furnishings). Operations were also broken down using *subsistence classes* but also *geography*. In addition, use was made of ten regional representatives whose advocacy function was described above.

A coordinator was used for programs, another for operations, and a third for

services. There was also a coordinator of regional representatives. The coordinators reported directly to the executive.

This design was modified slightly by CONASUPO executives so as to minimize the number of changes required in the current organization without sacrificing the essential features of the design. Although the resulting design of the organization as a whole has yet to be implemented, a 3D design of ICONSA, the production subsidiary of CONASUPO, has been completed and is in the process of implementation.

Universities

The output units of a university can be defined by types of degree awarded. Each of these, of course, are already called 'programs.' Research institutes or centers defined by subject matter, extracurricular programs, and athletic programs may also be included. Operating units consist primarily of academic departments coordinated by colleges. Service units include the library, computing center, printing services, financial and accounting departments, building and grounds, and so on.

In a university so organized academic programs receive income from tuition by vouchers turned in by students and cashed by the university's subsidizer, if it is subsidized. Program-benefit measures are some function of the number of degrees awarded. Therefore, they can be profit-and-loss centers. Research programs require other types of measure but they too can be profit-and-loss centers.

Academic programs buy teaching time and research programs buy research time from departments whose only income comes from sale of their services. If a program director is dissatisfied with a department's services he can buy services from another university or finance another department's development of the ability to provide the service he wants. Departments, on the other hand, can sell their services outside the university; for example to corporate educational programs or other universities.

This arrangement prevents control of a program by any one or combination of departments and it encourages interdisciplinary activities.

The presence of tenured faculty in departments whose time cannot be sold to programs creates a special problem. Unless tenure is modified it may be necessary for a department to be subsidized to cover its unsaleable and nondisposable faculty members. But at least this would be explicit and the executive knows where the academic deadwood is and how much it costs.

Charges made by a department for its services would be overheaded to cover its administrative costs and to provide it with some development funds.

Service units would also sell their services to programs and departments. They could also do so to the executive function.

The National Autonomous University of Mexico recently opened three new colleges on separate and dispersed campuses. Each of these has a design similar

to that described here. Financing of units, however, is a compromise between conventional budgeting of departments and the type of program budgeting described here. The University plans, however, to move toward this type of program budgeting. In the new colleges now being constructed and planned the University intends to take further steps in this direction.

ON DIVIDED RESPONSIBILITY

The MD design I have presented has something in common with what has been called 'matrix organizations'. But matrix organizations are normally two dimensional and lack many of the features of the design discussed above particularly the financial features. However, they do *appear* to have one deficiency in common: those in activity units appear to work for two bosses and this is generally taken to be undesirable. It is the most frequently cited deficiency of matrix organizations and is said to be the cause of what might be called 'occupational schizophrenia'.

The MD design creates no such problem although the matrix organization may. In an MD organization members of an activity unit whose activities have been purchased by a program director are related to that director no differently than if he were an external customer. They are responsible to the activity-unit director who, naturally, will use the program director's evaluation of his subordinates' performance in his evaluation of them. Those who head an activity unit's service group working for a program are in much the same position as are project directors in a construction or consulting firm. There is no ambiguity as to who their boss is. But they do have to deal with program directors as customers. There is nothing new or difficult about such a situation. As a director of research projects carried out by a university-based group for external sponsors, I am in exactly this position. The Dean of my college is my boss, not any of my sponsors. The sponsors, however, have considerable control over what I do and the Dean is well aware of how they evaluate my services.

Vol. 5
1977

MD DESIGN AND PROGRAM BUDGETING

The intersection and divergence of some of the ideas incorporated into the MD organizational design and the ideas contained in 'Program Budgeting' deserve some comment. Program Budgeting, as normally promoted and practiced, is only a way of preparing a budget for activities and programs; it is *not* associated with giving resources and choice to program units and with requiring activities to earn their way in an internal and external market place. In brief, Program Budgeting normally does not contain an *organizational* idea

and has no effect on organizational flexibility; it is a way of allocating resources to activities that provides greater assurance that programs will be carried out. In addition, it provides a more effective way than is usually available for determining the costs of programs. An MD design accomplishes all that a Program Budget does and much more.

CONCLUSION

The organizational design presented here is intended to increase organizational flexibility and responsiveness to changing internal and external conditions. It does so by dividing the organization into units whose survival depends on their ability to produce at a competitive price goods or services that are in demand. It creates a market place *within* an organization whether it be public or private, for-profit or not. Doing so increases responsiveness to both internal and external demand. Because units in an MD design are relatively independent of each other, they can be added, subtracted, or otherwise modified more easily than is the case in most conventional organizations. The performance measure of each unit is decoupled from that of any other. This makes it easier for the executive to evaluate and control them. Even the executive function can be evaluated separately. It is accountable for all of its decisions.

The design is also resistant to the development of bureaucracy. It is not possible for operating units or programs to be victimized by service units whose procedures become ends in themselves and obstructions to the pursuit of organizational objectives. Users, inside and outside the organization, control internal suppliers. Suppliers are never in control of users. The organization is end-, not means-oriented. Bureaucracy is characterized by the subordination of ends to means.

The MD design is not a panacea or a cure-all. Man cannot conceive of an organization that others are not capable of subverting. Although the MD design removes some of the stimulus to subversion found in many conventional organizations, it does not remove all of them. By itself the design gives no assurance of enriched and rewarding work at the lower levels of an organization. But the design makes it easier to apply other new ideas that are relevant to organizational development such as worker participation and self-determination.

The MD design is not the only way of increasing organizational flexibility and responsiveness to change, but thinking seriously about it increases flexibility and responsiveness of thinking about organizations. This, hopefully, will lead to new and superior organization designs.

ACKNOWLEDGEMENTS

In preparing this paper I have benefitted greatly from the constructive suggestions of my colleagues Thomas A. Cowan, Peter Davis, James R. Emshoff, Hasan Ozbekhan, Thomas L. Saaty, and Eric L. Trist.

REFERENCES

1. TOFFLER A. (1971) *Future Shock*. Bantam, New York.
2. SCHON D. (1971) *Beyond the Stable State*. Random House, New York.

ADDRESS FOR CORRESPONDENCE: *Russell L Ackoff Esq., University of Pennsylvania, The Wharton School, Busch Center, Vance Hall CS,* 3733 *Spruce Street, Philadelphia* 19174, *USA.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 6, 1977. Pergamon Press. Printed in Great Britain

The Decline of Venture Capital

N MEADE
Imperial College

(*Received May* 1977)

Venture capital is a means of financing the growth of companies with a technologically innovative product, the investor's objective is the realisation of capital gains on the sale of his shareholding. The history of venture capital in the US and the UK is examined. After attracting great interest in the late 1960s and early 1970s, activity in the venture capital field has declined significantly. The reasons for this decline are analysed and it is suggested that a recovery of interest is unlikely.

INTRODUCTION

VENTURE capital is a specialised form of high risk investment activity involving the provision of finance in the form of equity and loan to a company—the venture—to finance its growth. Generally the venture's objective will be to produce and market a technologically innovative product or service. The investor's long term goal is to realise a capital gain on the sale of equity in the venture. The details of the agreement between the investor and the venture will differ in terms of the management information and control required, but usually the equity participation of the venture capitalist will be less than 50%. The riskiness inherent in venture capital is caused by the innumerable pitfalls that may be encountered in the development from a small company with a promising technological innovation to a thriving concern with a saleable equity.

As a financial activity, venture capital attracted a great deal of publicity in the late 1960s and early 1970s, for example the Financial Times [9] was discussing 'venture capital fever' in December 1970. However, over 6 years later, attitudes have changed and the level of activity in venture capital is very low.

THE ORIGINS OF VENTURE CAPITAL

Venture capital became recognisable as a distinct form of investment activity during the 1930s in the US. At this stage, the investors were very wealthy families. A Rockefeller concern helped finance the formation of the Polaroid

Corporation in 1937 [19]. In 1945 activity increased and more private funds were set up with the purpose of investing venture capital.

A significant event in the history of venture capital was the founding of American Research Development (ARD) in 1946 in Boston by a group of bankers and academics from the Massachusetts Institute of Technology (MIT). ARD promoted a policy of involvement with the management of their ventures, fostering a strong relationship between the entrepreneur and the investor. Until 1960 ARD was the only venture capitalist with its shares publicly held [5].

Interest in venture capital increased significantly in the mid 1960s as the US economy was buoyant and it was seen as a new form of investment which was both profitable and exciting. Coincident with this surge of interest and primarily responsible for continuing interest was the flotation of Digital Equipment Corporation (DEC) in 1966.

ARD had invested \$61,400 for 78% of the equity of DEC in 1957 [5]. (The amounts of money mentioned by Dominguez [5] and New Enterprise Systems [19] do not correspond exactly, the figures mentioned are attributed individually). DEC manufactured small computers for scientific uses, exploiting a gap in the computer market left by the larger manufacturers. The company prospered and in 1966 after the public offering of DEC's shares, ARD's shareholding was worth \$53m [19]. DEC continued to prosper in January 1970, ARD owned 50% of the equity worth \$530m [19]. Share prices declined slightly and in December 1971 ARD owned 45% of the equity worth \$355m [5]. ARD made many other successful investments but none of the same magnitude as DEC.

Vol. 5
1977

A factor favouring investment of venture capital in the US was the Small Business Administration set up in 1958 by the US government to increase availability of capital to small businesses. The mechanism used was the Small Business Investment Company (SBIC), these companies were entitled to low interest loans from government with which they would finance small businesses. The rules under which SBICs operated were subject to fairly frequent government change. The SBIC programme was heavily promoted in 1960–61 and numbers of SBICs increased. However, it became apparent that the spirit of the original act was being misinterpreted or ignored; a high proportion of funds were being diverted to real estate development, funds were being re-loaned at higher rates rather than being used for equity purchase. In addition, many intending venture capitalists were inexperienced [5, 20]. Regulations were tightened and supervision was increased, as regulations became more restrictive some of the more successful venture capitalists left the programme. Despite attempts to revive the programme in 1968, investigations [19] showed that SBICs were not a significant force in venture capital and that if anything, their importance would decline further.

Venture capital activities are concentrated into a number of localities in the US. Campanella [2] identifies San Francisco and Los Angeles in the West, Houston and Dallas in the South, Minneapolis and Chicago in the North and

Boston and New York in the East. Boston is where ARD started and MIT is a major source of technological innovations in this area (Roberts [23] and Peters and Roberts [22] carried out surveys of innovations at MIT). Further evidence of the innovative nature of this area is the Route 128 complex near Boston. The complex is an industrial park consisting of companies exploiting advanced technology; Goodeve [12] identified government contracts and venture capital as the major sources of finance for these companies. Much of the technology exploited was developed originally for defence and space programme purposes. In 1972 MIT started its own small venture capital company in order to benefit financially from its own innovations (see Business Week [1]); unfortunately this company has missed many opportunities to participate in MIT-inspired successes including DEC.

Having dealt with the sources of venture capital and the sources of technologically innovative ventures, a third important facet of the venture capital process is the disposal of the investment in order to realise a capital gain.

In the US there is a large 'over the counter' (OTC) stock market. A share is bought or sold over the counter if the transaction did not take place using a formal stock exchange. The mechanism of the OTC market in the USA is fully described by Loll and Buckley [16]. The requirements that a company must fulfil to be quoted on the OTC market are significantly less stringent than those necessary for quotation on the registered stock exchanges such as the New York Stock Exchange (NYSE). It is the comparative ease with which companies can float their securities on a market which makes the OTC market relevant to venture capitalists. The venture capitalist may conveniently realise his investment in a venture by arranging to have the shares quoted on the OTC market. He need not necessarily sell his holding immediately, but the investment is a marketable security and thus far more liquid than prior to quotation.

According to Liles [15], venture capital in the US has become far more conservative in its investment approach in the mid 1970s. The sharp decline in prices of the shares of quoted ventures in 1972–1973, in line with the depressed economy, discouraged investors who had already seen the major proportion of investments fail for technical or managerial reasons. The actual proportion of venture capital investments that failed (i.e. where the venture capitalist lost money) is difficult to establish due to the understandable reluctance of venture capitalists to provide statistics. Campanella [2] estimates a failure rate of approx. 90%.

An analysis of the performance of publicly held venture capital companies [19] showed the median internal rate of return to be 8·9% p.a. (over 1960–1970). It was becoming increasingly apparent to investors that returns comparable to those gained in venture capital could be equalled with less risk. Fisher and Lorie [10] demonstrated a portfolio of shares quoted on the New York Stock Exchange would have yielded a median return of 8·6% p.a. over 1955–1965.

At the same time the Securities and Exchange Commission (SEC) were making

the sustenance of venture capital operations difficult. A rule governing the sale of unregistered shares requires an owner of more than 1% of a company's equity to hold it for at least 2 years and then sell no more than 1% of the equity every 6 months [5]. This rule obviously prohibits quick disposal of a sizeable equity holding in a company. SEC restrictions were a contributing factor to the acquisition of ARD by Textron, a conglomerate in 1972. One problem experienced by ARD was its inability to compete for executive staff because it was constrained by the Investment Company Act from offering incentives such as stock options to its employees. The act did not apply to ARD in its new situation.

Venture capital in the US has returned to the province of the extremely rich and the large corporations (see Forbes [11]). The costs of starting businesses have increased drastically, but the major reason for the lack of enthusiasm for venture capital appears to be the less sanguine view taken by the stock markets. The price/earnings ratios on new ventures are too low. In order to raise cash for expansion by selling equity the holdings of the venture capitalist and the entrepreneur would be too diluted to make the exercise worthwhile. As pointed out in Forbes [11] well established companies are more profitable and less risky investments. The new issue market reflects this conservatism, in 1969, 698 issues of shares in small companies were made in the US, in 1974 the comparable figure was 9.

Vol. 5
1977

FINANCE FOR INNOVATION IN THE UK

The consideration of finance for innovation can be divided into two general cases depending on the nature and intentions of the innovators. If the innovator seeks commercial exploitation of his work via a company then he has the problems associated with a 'small business'. However, he will also be described as seeking 'finance for innovation', this title includes the licensing of larger companies to exploit innovation and other non-entrepreneurial ways of exploiting innovation. Both phrases 'small businesses' and 'finance for innovation' have come periodically into vogue over the last 50 years.

In 1971 the Committee of Inquiry on Small Firms [3], chairman Mr. JE Bolton, published its report. This report is the most recent of a series. In 1931 the Macmillan Committee on Finance and Industry reported that there was a lack of provision of long term finance for amounts less than £200,000, which was hampering the development of small businesses. This phenomenon became known as the 'Macmillan Gap'. A number of companies were formed to help bridge the Gap, a notable example being Charterhouse Industrial Finance Ltd. However, those companies active at the time were able to satisfy only a small proportion of the demand for finance [3]. In 1945 a cabinet committee recommended an increase in the resources being used to close the Macmillan Gap. This recommendation led to the formation of the Industrial and Commercial

Finance Corporation (ICFC), financed by the Bank of England, the Clearing and Scottish Banks. ICFC was to supply equity and loan finance in the range £5000–£200,000, the upper limit being increased by the Radcliffe committee in 1959 [3, 14]. One of the central conclusions of the Committee of Inquiry on Small Firms was that there was no 'Bolton Gap'.

As a consequence of a desire to formalise the innovation process in the UK, the National Research Development Corporation (NRDC) was created by the Development of Inventions Act in 1948. NRDC is government financed through the Department of Industry and designed to act in the public interest acquiring and granting the rights to the development and exploitation of inventions. In its early years NRDC concentrated on co-operation with universities and government institutions. However, in 1965 a deliberate effort was made to increase the number of projects jointly supported with industry (see NRDC report 1965–1966 [18]). In 1968–1969, 121 new projects were started, this being a peak figure, whilst in 1975–1976 only 82 new projects were started. As a general rule, NRDC derives its return from a levy on sales revenue, however it has subsidiary companies (shareholding exceeds 50%) and associated companies (shareholding is between 10% and 50%). These companies can legitimately be termed NRDC's venture capital portfolio and in 1975 there were six subsidiary companies and twenty-five associated companies valued at £191,000 and £2·43m respectively. The income from these investments is negligible, over 75% of income comes from licensing agreements (£9·04m in 1974/5). The highest earning licences relate to an antibiotic, namely Cephalosporin, and insecticides and weedkillers.

The Radcliffe Committee, reporting on the Working of the Monetary System in 1959, recommended that a corporation be created to facilitate the commercial exploitation of technical innovations. This led to the foundation of Technical Development Capital (TDC) in 1962, which became a subsidiary of ICFC in 1966 [25]. TDC remains active in venture capital with a large number of ventures in its portfolio.

Having dealt with the officially instigated institutions, let us move on to the private sector. In the mid 1960s the attractions of venture capital were being communicated from the US. As Holborn and Edwards [14] point out, quoted shares were fully valued and investments in private companies were the main opportunities for investment growth. Other companies were set up to join TDC as venture capitalists, Midland Montagu Industrial Finance was formed by a clearing bank and a merchant bank and Small Business Capital Fund was formed by an insurance company and a market research group. Many merchant banks set up venture capital divisions or acted as agents for the provision of venture capital.

In addition to enthusiasm being communicated across the Atlantic, expertise was also being exported. European Enterprises Development (EED) was set up in Luxembourg in 1964 with equity from British, European and American banks

and ARD who also supplied the president. EED was designed to act as a venture capitalist to the whole of Europe. In 1972 (the year of the last available annual report), EED had thirty ventures spread through Europe, fourteen in France, three in the UK [8].

DISILLUSIONMENT WITH VENTURE CAPITAL IN THE UK

Generally venture capital receives publicity for two reasons; firstly when a new company is financed and shows both technological and commercial promise, secondly when an investment has failed and money has been lost. Unfortunately, it is very rare for any successful disposals of ventures to be reported in the UK.

The catalogue of unsuccessful venture capital operations is long and only a few examples will be mentioned.

Spiegelberg [24] describes the history of Spey Investments. The company was formed in 1965 to invest institutional funds in small companies. In excess of £50m was provided by institutions such as the pension funds of ICI, Unilever and Barclays Bank. After a period of great activity, it became apparent that some of the investments were showing large losses. The fund went into a steep decline, the founder resigned and the Unilever Pension Fund withdrew and the institutions made a significant net loss.

Vol. 5
1977

The Economists Advisory Group (EAG) in their report for the Bolton Committee [6] cite a case of a venture capital investment company set up in 1966 by a leading merchant bank and an insurance company. The company was to invest in any area, including innovative products, where there was a prospect of growth. The company found great difficulty in finding worthwhile potential investments and after 4 years only three investments had been made and winding up of the company commenced. No information is given about the progress of the investments.

In 1969, Miles Roman, a company owned by Kleinwort Benson formed a venture called Autonomics. The company was to provide easy-to-use computer terminals to small organisations. Loans were obtained from Ferranti Ltd. (one of the equipment suppliers) and NRDC (see NRDC report 1970–1971 [18]). The first installations were made during 1970. The company began to experience severe cash flow problems in 1971. In August 1971 Kleinwort Benson made a loan to the company and in October 1971 NRDC appointed a receiver. The receiver reported in March 1973 that unsecured creditors of Autonomics would not be paid and there would be a shortfall of £980,000 to debenture holders (see Times [27]).

Not particularly innovative technologically, but instructive as an example of a failed venture is the case of Carousel Parties (Lingerie) [26]. This was a joint

investment by a private individual and Bankers Trust International (BTI) made in 1970. Its business was the sale of lingerie at specially arranged parties. The investors provided £600 out of £1000 equity and £60,000 in loan capital. Stock control and cash flow problems quickly led to additional investment being required. After providing £25,000 in September 1971, BTI appointed a receiver in December. A BTI director implied that the company was unlikely to experiment further in venture capital.

A number of companies mentioned in various articles as investors of venture capital ceased activity fairly quickly after the initial publicity.

The most notable recent event (October 1976) is the placing of EED, one of the foremost European venture capitalists, into a state of 'gestion contrôlée' by its shareholders [28]. This is a state of suspended animation during which the group is protected against claims on its assets while the courts decide whether the group is able to proceed on a sound financial basis.

DISINCENTIVES TO INVESTING VENTURE CAPITAL IN THE UK

Long term interest rates have steadily increased since 1970 from about 10% to about 16% at the end of 1976. This increase has made investment in venture capital steadily less attractive. There is little evidence to show that returns on venture capital will compare favourably with interest rates of 10–12%. The NES [19] figures quoted earlier for US venture capitalists are not encouraging; only 6 out of 15 earned more than 10% whilst only ARD earned more than 18·5%. There is no incentive to take the risks inherent in venture capital unless very high returns are very likely, an uncommon situation.

A further dampening effect on venture capitalists is the increased price of entry into many of the markets that ventures exploit. Since the mid 1960s, the computer industry has progressed technologically and commercially. As a consequence, the difficulty and cost of making significant improvements or innovations have both increased. Given the innovative idea, the basis of a venture, far more funds are necessary to demonstrate its viability than were required in 1965. An example of this contrast is quoted in Forbes [11]. In the mid 1960s Scientific Data Systems delivered their first computer after 9 months and $1m expenditure. To achieve the same task, Amdahl, a newer computing venture, took from 1971 to 1975 and over $100m. Since a large proportion of venture capital has been invested in the computer industry, the increased price of market entry will be a significant discouragement to further investment.

In addition to the prevalent economic climate, there are more fundamental structural reasons why venture capital has ceased to flourish and indeed is unlikely ever to flourish in the current environment. The early experience with well publicised failures increased caution on the part of investors and it became

apparent that there was no abundance of worthwhile ventures. The EAG [6] case study is an example of a venture capitalist winding up due to a lack of suitable investments. It would be interesting to compare the rates of production of potential ventures in the UK and the US but only very limited data are available.

Gibbons and Watkins [12] compared the propensities to innovate of UK and US scientists and found that although US scientists were hampered mainly by lack of time, UK scientists cited lack of finance. However, the UK scientists were unaware of many of the available financial sources and Gibbons and Watkins comment "if anything there may be a shortage of people coming forward with suitable commercially oriented ideas".

The results of a sample survey carried out by the EAG [7], to discover the problems encountered by small firms in raising finance, confirmed the opinion that there was no serious deficiency in the supply of finance. EAG felt, on the contrary, that the supply of finance might be too generous, and that some finance was being used inefficiently. Commenting on the EAG's findings, the Bolton Committee noted that this exercise had illustrated the "extreme difficulty in pinning down genuine specimens of that all-important type—the small firm which needs finance, deserves it, has made serious efforts to raise it and has been turned down".

Vol. 5
1977

One factor that contributes to the disillusionment in venture capital is the impatience of investors who want to see returns on their investments as quickly as possible. If the time from investment in a venture to its sale is between 4 and 10 years, then it is obvious that it will take a portfolio of ventures a minimum of 10 years to reach an equilibrium. That is a situation where the pattern of investments and sales is such that one year's earnings is a useful indicator of the following year's. Using a simple stochastic model of a venture capitalist's portfolio with the average life of a venture equal to 4 years, Meade [17] demonstrated that a stable earnings pattern developed only after about 12 years. This casts doubt on the value of the figures quoted by NES [19] on the returns of US venture capialists as most of them had been in business 8 or 9 years.

A probably more fundamental problem than those already mentioned is that the machinery for raising company finance is inconsistent with the aims of the venture capitalist. If a venture capitalist is going to remain in business, he needs to be able to dispose of his investments. In the US this procedure is facilitated by the 'Over the Counter' (OTC) market in shares, which has less stringent rules about the securities traded than the registered stock exchanges. According to Campanella [2] a number of virtually untried ventures offered stock on the OTC market in the "speculative excesses of the 1966–68 era". When normality returned, the prices were generally well below their initial values. Aside from this exploitation of the market, the central point remains that the existence of a strong secondary market is vitally important for the success of venture capital operations. As Campanella points out, without the market, funds would become

locked into ventures and thus the number of new ventures being financed would drop.

In the UK there is no parallel to the OTC market and it is impossible to offer the shares of small companies publicly. Davis [4] indicates that companies using the UK new issue market need not only to be big but to have a consistent record of profit growth. Peretz [21] mentions the idea of the 'equity bank' but dismisses it as it is built on a false premise that the problem is shortage of finance and in any case ICFC basically fulfills the rôle.

The problem is to make holdings in small companies more liquid so they can be valued and traded more easily. The Bolton Committee [3] considered an OTC market to be of no advantage in the UK since the relevant companies would be too small. It also rejected the idea of a unit trust specalising in small private companies because of high administration costs and 'insuperable' problems of share evaluation.

Another alternative open to the venture capitalist is the sale of a venture to larger companies with interests related to the venture. With a necessarily small number of customers, the difficulties of arranging a sale may be insurmountable. Some companies build up a portfolio of ventures with no intention of sale, but with the objective of forming a group of companies in the same industry. It is debatable whether this process can justifiably be called venture capital.

Since any kind of suitable secondary market seems to be ruled out, it seems unlikely that venture capital will flourish outside the structured system of the government run or government inspired institutions—NRDC and TDC respectively. This is not to say that other financial institutions no longer indulge in venture capital, but they also face the problems described in this paper. It is no longer a glamorous activity and the expectations of results are far more pessimistic than in the early 1970s.

CONCLUSION

The rise to prominence of venture capital in the US and the UK and its subsequent decline have been examined.

The structure of financial markets in the US are shown to be significantly more favourable to the venture capitalist than in the UK, although increasingly strict regulations are having a dampening effect. The economic climate coupled with the maturity of many of the industries that fostered ventures in the past have made venture capital unattractive to both investor and entrepreneur. Combined with a record when failures outnumber successes by a considerable majority, the factors described coalesce into a body of evidence suggesting that a recovery of interest in venture capital is unlikely, especially in the UK.

REFERENCES

1. *Business Week* (1972) MIT aims to grow a new route 128. **2236**, 72–73.
2. CAMPANELLA FB, KELLEY AJ and MCKIERNAN J (1971) *Venture Capital—A Guidebook for New Enterprises*. The Management Institute, Boston College, Massachusetts.
3. COMMITTEE OF INQUIRY ON SMALL FIRMS (1971) *Small Firms* (Chairman JE Bolton) HMSO London.
4. DAVIS EW (1976) Finance and industry—the debate continues. *Banker* **126**(603), 473–476.
5. DOMINGUEZ JR (1974) *Venture Capital*. Lexington Books, Massachusetts.
6. ECONOMISTS ADVISORY GROUP (1971) *Financial Facilities for Small Firms*. Committee of Inquiry on Small Firms Research Report No. 4, HMSO London.
7. ECONOMISTS ADVISORY GROUP (1971) *Problems of the Small Firm in Raising External Finance: the Results of a Sample Survey*. Committee of Inquiry on Small Firms Research Report No. 5, HMSO London.
8. EUROPEAN ENTERPRISES DEVELOPMENT (1972) EED Annual Report, Luxembourg.
9. *Financial Times* (1970) Venture capital and Europe (p. 16, 11 December).
10. FISHER L and LOPIE JH (1968) Rates of return on investments in common stock. *Jl Bus. Univ. Chicago* (July) p. 1–26.
11. *Forbes* (1975) (Semi-monthly business journal published in New York). Various articles in **116**(10), 22–34.
12. GIBBONS M and WATKINS DS (1970) Innovation and the small firm. *R&D Mgmt* **1**(1), 10–13.
13. GOODEVE C (1967) A Route 128 for Britain? *New Scientist* **47**, 346–348.
14. HOLBORN PRM and EDWARDS ECN (1971) Development of the UK venture capital market. *Banker* **121**, 488–492.
15. LILES PR (1974) *New Business Ventures and the Entrepreneur*. Irwin, Homewood, IL.
16. LOLL LM and BUCKLEY JG (1961) *The Over the Counter Securities Markets—A Review Guide*. Prentice-Hall, New Jersey.
17. MEADE N (1975) Applications of Management Science in Assessment of Venture Capital. Ph.D. Thesis, University of London.
18. NATIONAL RESEARCH DEVELOPMENT CORPORATION (1965–1976) Annual Report and Accounts for years 1964–65 to 1975–76. HMSO London.
19. NEW ENTERPRISE SYSTEMS (1970) Venture Capital in the United States: An Analysis. NES Inc. Massachussets.
20. NOONE CM and RUBEL SM (1970) *SBICs: Pioneers in Organised Venture Capital*. Capital, Chicago.
21. PERETZ D (1976) Finance for investment—issues and non issues. *Banker* **126**(602), 365–372.
22. PETERS DH and ROBERTS EB (1969) Unutilised Ideas in University Laboratories. *Acad. Mgmt Jl* **12**, 179–191.
23. ROBERTS EB (1970) How to succeed in a new technology enterprise. *Technol. Rev.* **73**(2), 23–27.
24. SPIEGELBERG R (1973) *The City-Power without Accountability*. Quartet, London.
25. STEVENS HAR (1969) Sources of finance for innovation. *New Technol.* **31**, 3–6.
26. *Sunday Times* (1972) Bankers Trust gets caught in a spin on Carousel. **56** (16 January).
27. *The Times* (1973) Autonomics receiver report. 21 (27 March).
28. *The Times* (1976) Venture capital group meets a cash crisis. 19 (21 October).

ADDRESS FOR CORRESPONDENCE: *Dr Nigel Meade, Dept of Management Science, Imperial College, Exhibition Road, London SW7 2BX, UK.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 6, 1977. Pergamon Press. Printed in Great Britain

A Profitability Model for Tactical Planning

S EILON

GP COSMETATOS

Imperial College, London

Received June 1977)

Simple models reflecting changes in aggregate company behaviour in response to changes imposed by management decisions and/or factors outside the direct control of management provide useful tools for tactical planning purposes.

One such model is discussed in this paper. The model enables management to assess explicitly and quantitatively the effect on the rate of return on capital employed of primary causes of change in any of the key variables affecting company performance, such as the unit price, the unit cost, the output level and the capital employed. Simple functional relationships are proposed between the major variables in the model, such as the level of output and the unit cost, the unit price and the volume of sales, the working capital requirements and the general level of the firm's activity, etc. and these are based on analytical considerations, empirical evidence and a series of assumptions. These relationships give rise to a set of parameters which are incorporated in the model and describe the operational (cost structure), financial (capital structure) and marketing (elasticity of demand) position of the company under study.

Numerical results obtained from the model seem to suggest that the sensitivity of the rate of return to some of the parameters and underlying assumptions is relatively small, implying that the inter-dependence between some of the variables can be ignored, thus simplifying the problems associated with data collection and model evaluation.

Both the model and the methodology of analysis employed are flexible enough to allow: (1) application of the model to alternative sets of functional relationships between variables and/or underlying assumptions; (2) further decomposition of the model to include additional variables depending on the particular objectives of the study; (3) modification of the model to explore the effects of management decisions or those of changes in external factors on productive unit and product profitability; and (4) analysis of models reflecting the effects of changes which are imposed on the system on measures of business performance other than the rate of return on capital employed.

INTRODUCTION

IN ITS simplest form, Gold's model for managerial control ratios [5] can be expressed as

$$r = \frac{(p-c)V}{I} \tag{1}$$

where r = rate of return on total investment = (profit before tax)/(total investment);
p = unit price for the output;
c = unit cost for the output;
V = level of sales assumed equal to the vel of output, measured in physical terms;
I = total investment.

Since the rate of return r is widely regarded as a key measure of business performance, it is desirable to examine the possible effect of management decisions on this measure. As we see from (1) these decisions, coupled with factors outside the direct control of management, relate to the four elements on which r depends, namely decisions that affect the unit price, the unit cost, the output level and the total investment. If we assume that these elements are subject to change, so that

p changes to $p+\delta p = p(1+p^*)$;
c changes to $c+\delta c = c(1+c^*)$;
V changes to $V+\delta V = V(1+V^*)$;
I changes to $I+\delta I = I(1+I^*)$,

where δp, δc, δV and δI are the incremental changes and p^*, c^*, V^* and I^* denote the relative changes (i.e. $p^* = \delta p/p$, etc.), then the effect on r is that it becomes $r+\delta r$. It is not difficult to show that the relative change $r^*(= \delta r/r)$ is obtained by

$$r^* = \frac{1}{1+I^*}\left\{\frac{1+V^*}{\alpha}[p^*-(1-\alpha)c^*]+V^*-I^*\right\} \qquad (2)$$

with $\alpha = (p-c)/p$ denoting the original profit margin (prior to the changes in p and c).

The expression (2) provides an analytical tool for assessing the effects on r of changes in its constituent elements. It should be noted, however, that these elements are often inter-related and that any change in the system, even when it is primarily directed towards one element, has an effect on the others; it is only when all the relative changes have been established that their combined effect on r can be determined by (2). It would be tempting, for example, to examine the effect of a change in the selling price on the rate of return by setting $c^* = V^* = I^* = 0$ and to conclude that $r^* = p^*/\alpha$. Such a result would only be valid if the other relative changes have indeed been ascertained to be zero, but in reality a change in price may have an effect on working capital (and hence on I, which in turn may have an effect on interest charges and hence on c) and on demand (which may have a further effect on c). In using (2) it is, therefore, necessary to identify not only the primary changes that may have occurred, but also the secondary changes, the values of which need to be determined by empirical evidence and approximate functional relationships between p, c, V and I.

Vol. 5
1977

This paper is confined to the analysis of the short term effects on r of primary and secondary causes of change. Assumptions regarding the functional relationships between variables are made and the sensitivity of r to these assumptions examined. The model in (2) is shown to provide a useful means for predicting how a given system is likely to behave as a result of a set of changes which are imposed on the system both by outside factors and by management decisions.

ASSUMPTIONS AND FUNCTIONAL RELATIONSHIPS

It is conventional to assume that the unit cost takes the form (see, for example [4])

$$c = s + \frac{F}{V} \tag{3}$$

where s = the variable cost per unit of output;
F = the fixed costs;
V = output.

The interest charges on bank loans and overdrafts (to finance part of the working capital) are subsumed in s or in F. To take account of the increasing concern in many industries regarding the effect of interest charges on business performance we propose to add a term which explicitly expresses the cost resulting from these interest charges; thus,

$$c = s + \frac{F}{V} + \frac{J}{V} \tag{4}$$

where $J = jB$ = the interest paid on the total sum B (being the loans and overdrafts) at an average rate of j (which is then excluded from the definitions of s and/or F).[1]

If we denote the cost proportions in (4) as f_1, f_2 and f_3, so that $s = f_1 c$, $F/V = f_2 c$ and $J/V = f_3 c$ (note that $f_1 + f_2 + f_3 = 1$), it can then be shown that changes in the elements on the right-hand side of (4) will lead to

$$c^* = f_1 s^* + f_2 \left(\frac{F}{V}\right)^* + f_3 \left(\frac{J}{V}\right)^*$$

or

$$c^* = f_1 s^* + \frac{1}{1+V^*}[f_2 F^* + f_3 J^* - (1 - f_1) V^*]. \tag{5}$$

In the special case (3), when a simpler version of (4) is considered (by putting $J/V = 0$) and when it is assumed that both the unit variable cost s and the fixed costs F remain unchanged (so that $s^* = F^* = J^* = f_3 = 0$), equation (5) is reduced to

$$c^* = -\frac{f_2 V^*}{1+V^*}$$

which is the expression cited in [4].

[1] It is possible, of course, to disaggregate the sum B into $B = B_1 + B_2 + \ldots$, where the separate loans B_1, $B_2 \ldots$ attract interest rates j_1, $j_2 \ldots$, so that the total interest paid is $J = j_1 B_1 + j_2 B_2 + \ldots$

It is possible to expand (5) further by decomposing J^* and substituting (following the method shown in [2])

$$J^* = j^* + B^* + j^*B^* \tag{6}$$

It is also noteworthy from (5) that in the event that

$$\frac{f_3}{f_2} = \frac{F^* - V^*}{V^* - J^*}$$

the changes in the fixed costs, in the interest charge and in the volume annul each other, the expression in the square brackets becomes zero and (5) is reduced to

$$c^* = f_1 s^*.$$

Turning to the total investment, we assume that it can be expressed as

$$I = I_W + I_F = A + B + I_F$$

where I_W = working capital, of which the amount B is financed from bank loans and overdrafts, and we denote $w = I_W/I$ and $\lambda = B/I_W$, so that the interest charge is $J = \lambda j I_W$

I_F = fixed investment = $(1-w)I$.

Thus, the relative change in total investment is

$$I^* = wI_w^* + (1-w)I_r^* = w[(1-\lambda)A^* + \lambda B^*] + (1-w)I_r^* \tag{7}$$

It is assumed that the cost of capital B exceeds the cost of capital A so that any change in working capital requirements affects the borrowing only (i.e. if $I_W^* \neq 0$ then $B^* \neq 0$ but $A^* = 0$); to simplify the analysis it is also assumed that there is no change in fixed investment in the short term ($I_F^* = 0$). Hence we have

$$I_w^* = \lambda B^* \tag{8}$$

and

$$I^* = wI_w^* = \lambda w B^*. \tag{9}$$

An analysis of the effects of various changes in the system on the key factors in equation (1), and hence on the rate of return, may proceed on the basis of certain functional relationships between the variables involved. We state below our assumptions regarding these relationships, but it should be emphasised that the methodology employed in this paper can be easily applied to alternative sets of assumptions, also that the degree to which the model may be further decomposed to include more variables is merely a matter of convenience and emphasis.

The analysis of this paper is confined to five cases, each involving a simple primary cause of change in one of the variables s, F, j, V and p respectively. The assumptions relating to the effects of each primary change on the other variables involved in our model are summarised in Table 1 (where 0 denotes no effect and + indicates a possible effect).

In case 1 it is assumed that a change in the unit variable cost has no effect on the fixed costs F, the interest rate j, the volume V or the price p; but the unit cost

Vol. 5
1977

TABLE 1. ASSUMPTIONS ON THE EFFECTS OF PRIMARY CAUSES

Case	Primary cause	Effect on s	F	j	V	p	c	I_W	B	J	Assumptions
1	s		0	0	0	0	+	0	0	0	(10)
								+	+	+	(11)
2	F	0		0	0	0	+	0	0	0	(10)
								+	+	+	(11)
3	j	0	0		0	0	+	0	0	+	(10)
								+	+	+	(11)
4	V	0	0	0		0	+	+	+	+	(10), (11)
5	p	0	0	0	+	+		+	+	+	(10), (11)

c is obviously affected by (4) and the question arises regarding the possible effect of a change in unit cost on the working capital requirements. No single relationship between working capital and the level of activity in the firm is universally accepted (see, for example [6]) and we may proceed to explore two possible assumptions:

$$I_w^* = g(pV)^* = g(p^* + V^* + p^*V^*) \tag{10}$$

or

$$I_w^* = h(cV)^* = h(c^* + V^* + c^*V^*) \tag{11}$$

In (10) it is suggested that a change in the working capital is proportional to a change in the turnover pV, whereas in (11) it is proportional to a change in the total cost cV. In either case the assumption is that for relatively small changes in the values of the variables, expressions such as (10) and (11) are realistic and that an experienced financial controller can estimate the value of the proportionality coefficient g or h for the current level of activity of his company. Now, if (10) applies, then for case 1 the value of pV remains unchanged, hence I_W and therefore B and J are unaffected, but if (11) operates, all three variables are subject to change; the two alternatives are shown in the Table.

A more general assumption is to combine (10) and (11) into

$$I_w^* = g(pV)^* + h(cV)^* \tag{12}$$

which allows both the effects of the turnover pV and the total cost cV to be accounted for. The alternative assumptions in (10) and (11) can then be derived as special cases with $h = 0$ and $g = 0$ respectively.

In case 2 in Table 1 it is assumed that a change in F has no effect on s, j, V or p, but again the unit cost c is affected and the consequences for I_W and hence B and J depend on whether (10) or (11) is adopted, as in case 1. In case 3 we consider a change in the interest j and assume that it does not lead to changes[2] in

[2] It is not suggested that such an assumption is always valid: a change in j may well be linked to changes in the economic climate and these are likely to affect demand and price level. It is not difficult to expand the analysis in this paper to account for such changes when their magnitude or the relationships between them are known.

s, F, V or p, but that it has an immediate effect on J (even if the borrowings B remain unchanged); the effect on I_W and B again depends on (10) or (11).

In case 4 we assume that a change in demand resulting from external forces has no effect on s, F, j or p, but that it does affect the other variables. Finally, in case 5 we assume that a change in the unit price does not affect the cost components s and F nor the interest rate j, but that it does affect the level of demand V and hence all other variables.

CASES 1, 2 AND 3: A CHANGE IN s AND/OR F AND/OR j

As it is assumed that s, F and j are independent of each other, the first three cases in Table 1 may be considered together. The pattern of their respective effects on the other variables is similar in kind: it is assumed in Table 1 that changes in s, F and j have no effect on V or p. As for their effect on the unit cost c and on the rate of return r, it depends whether assumption (10) or (11) is adopted. In the case of (10), we conclude from the Table and from (9) that the borrowing requirement B is unaffected and that $I_W^* = I^* = 0$. The interest charge J for cases 1 and 2 is also unchanged, but in case 3 it follows from (6) that $J^* = j^*$. Hence the primary changes s^*, F^* and j^* have a direct effect on the unit cost. To simplify the notation in subsequent sections of this paper, the resulting direct relative change in c, as calculated from (5) is denoted by c_o^* and is given by

$$c_o^* = f_1 s^* + f_2 F^* + f_3 j^*. \tag{13}$$

c_o^* is the only factor that affects the rate of return r as shown in Fig. 1, and the consequent relative change r^* is obtained from (2)

$$r^* = -\frac{1-a}{a} c_o^* \tag{14}$$

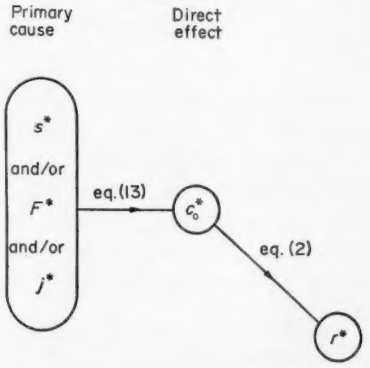


FIG. 1. *The effect of* s* *and/or* F* *and/or* j* *for assumption* (10).

Vol. 5
1977

Case 1 in Table 1 is now obtained as a special case where $F^* = j^* = 0$, case 2 where $s^* = j^* = 0$ and case 3 where $s^* = F^* = 0$. Equations (13) and (14) thus determine the relative changes in unit cost and in the return on investment, if assumption (10) holds.

For the more general assumption (12) (which incorporates (11)), the relative direct incremental change in unit cost c_o^* does have an effect on the working capital, which in turn affects the total investment as well as the borrowings B. A relative change in B then affects the interest cost J, which has a further effect on the unit cost, and the pattern of transmitted effects then repeats itself, as shown in Fig. 2. The final values of c^*, I^* and r^* may therefore be derived from a converging iterative process. To obtain the secondary effects on c we start by putting $p^* = V^* = 0$ in (12), so that $I_W^* = hc_o^*$. Now, from (8) we have $B^* = (h/\lambda)c_o^*$ and the expression in (6) then gives $J^* = j^* + (1 + j^*)(h/\lambda)c_o^*$; hence if we denote

$$u = (1 + j^*)f_3 h/\lambda \tag{15}$$

we get from (5) with $V^* = 0$

$$c^* = c_o^*(1 + u)$$

where uc_o^* is the first order secondary effect on the unit cost, triggered off by the change c_o^*, which in itself is directly caused by the primary relative changes s^*, F^* and/or j^* (see Fig. 2). The new value of c^* has a further secondary effect on the unit cost, and we can proceed as before to obtain $I_W^* = hc^*$, then $B^* = (h/\lambda)c^*$ and $J^* = j^* + (1 + j^*)(h/\lambda)c^*$, except that now c^* is taken as $c_o^*(1 + u)$. Using (5) we now get a new value $c^* = c_o^*(1 + u + u^2)$. It is not difficult to verify that the limiting value of c^* is thus given by

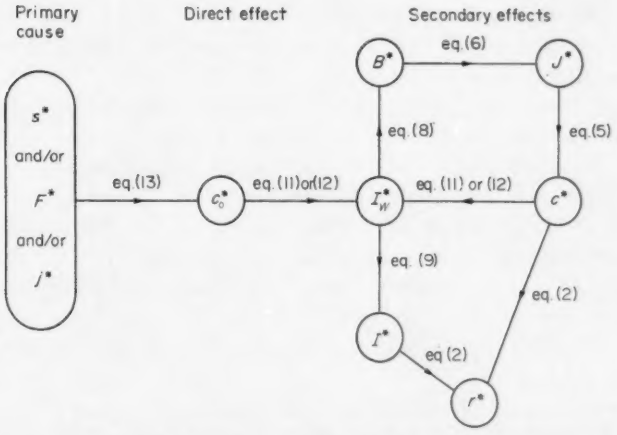


FIG. 2. *The effect of* s* *and/or* F* *and/or* j* *for assumption* (11) *or* (12).

$$c^* = c_o^*(1 + u + u^2 + u^3 + \ldots)$$
$$= c_o^*/(1 - u) \quad \text{for } u < 1 \tag{16}$$

Alternatively, (16) can be obtained directly from (5)[3].

The convergence of c^* in (16) depends on the condition $u < 1$. The value of u can be rewritten by substituting $f_3 = jB/(cV)$ and $\lambda = B/I_W$ into the definition (15), so that

$$f_3/\lambda = jH \tag{17}$$

and

$$u = j(1 + j^*)hH \tag{18}$$

where $H = I_W/(cV)$ is the ratio of working capital to total cost. Since both h and H are generally expected to be less than 1, equation (18) suggests that even with relatively high rates of interest the value of u is sufficiently small to justify the rapid convergence of the iterative process described in Fig. 2 (for example, if the new rate of interest $j(1 + j^*)$ is 15% and if $h = H = 0{\cdot}5$ then $u = 0{\cdot}038$, and even if $h = H = 1$ then $u = 0{\cdot}15$).

The limiting value of I^* can now be obtained from (9) and (11) as $I^* = wI_W^* = whc^*$, so that when I^* is substituted into (2)

$$r^* = -\frac{c^*}{1 + whc^*}\left(\frac{1 - \alpha}{\alpha} + wh\right) \tag{19}$$

where c^* is given by (16).

Thus, equations (16) and (19), where c_o^* is given in (13), represent general results for each of the cases 1–3 in Table 1, or for any combinations of these cases. If assumption (10) is appropriate, namely that a relative change in working capital is proportional to a relative change in turnover, then $h = 0$, and hence $u = 0$; when these are substituted in (16) we obtain $c^* = c_o^*$ (as depicted in Fig. 1) and (14) becomes a special case of (19).

Numerical values of r^* in percentage terms for $j(1 + j^*) = 0{\cdot}15$, $c_o^* = 0{\cdot}01$, $w = 0{\cdot}5$ and for $\alpha = 0{\cdot}20$, $0{\cdot}10$, $0{\cdot}05$ are presented in Table 2, from which we may draw the following conclusions:

(a) Small changes in s, F or j leading to a direct change in unit cost of just 1% have a marked effect on the rate of return r.

(b) The sensitivity of r to h and H is rather low, and particularly for low values of H the secondary effects on unit costs can be ignored by assuming $h = 0$ and adopting (14) as a good approximation for (19). Furthermore, an error in estimating h can have a relatively small effect on the result in computing r^*, and to a lesser degree this conclusion holds for an error in estimating H.

(c) $h = 0$ in (11) implies that the working capital requirements are insensitive

[3] For $V^* = 0$ equation (5) is reduced to (13). If we substitute into it $J^* = j^* + (1 + j^*)B^* = j^* + (1 + j^*)I_w^*/\lambda = j^* + (1 + j^*)(h/\lambda)c^*$ as obtained from (6), (8) and (11), and solving for c^* we get the expression in (16).

Vol. 5
1977

TABLE 2. VALUES OF r^* (IN %) BY EQUATION (19) FOR $w = 0{\cdot}5$, $c_o^* = 0{\cdot}01$ AND $j(1+j^*) = 0{\cdot}15$

(i) $\alpha = 0{\cdot}20$

H \ h	0	0·2	0·4	0·6	0·8	1·0
0·2	−4·00	−4·12	−4·24	−4·37	−4·49	−4·62
0·4	−4·00	−4·15	−4·29	−4·45	−4·60	−4·76
0·6	−4·00	−4·17	−4·35	−4·53	−4·72	−4·92
0·8	−4·00	−4·20	−4·40	−4·62	−4·85	−5·09
1·0	−4·00	−4·22	−4·44	−4·71	−4·98	−5·26

(ii) $\alpha = 0{\cdot}10$

H \ h	0	0·2	0·4	0·6	0·8	1·0
0·2	−9·00	−9·15	−9·29	−9·44	−9·59	−9·74
0·4	−9·00	−9·20	−9·41	−9·62	−9·83	−10·05
0·6	−9·00	−9·26	−9·52	−9·80	−10·09	−10·38
0·8	−9·00	−9·31	−9·64	−9·99	−10·35	−10·73
1·0	−9·00	−9·37	−9·77	−10·19	−10·63	−11·11

(iii) $\alpha = 0{\cdot}05$

H \ h	0	0·2	0·4	0·6	0·8	1·0
0·2	−19·00	−19·20	−19·39	−19·59	−19·80	−20·00
0·4	−19·00	−19·31	−19·63	−19·96	−20·29	−20·63
0·6	−19·00	−19·43	−19·88	−20·34	−20·82	−21·31
0·8	−19·00	−19·55	−20·13	−20·73	−21·37	−22·03
1·0	−19·00	−19·67	−20·38	−21·14	−21·95	−22·81

to small changes in total costs. In such a case, $c^* = c_o^*$ and $I_W^* = I^* = 0$, hence (19) is reduced to (14). Therefore, the values in the table corresponding to $h = 0$ measure the effect on r of the direct change in unit cost; the values in the other columns measure the effect on r of both the direct and secondary changes in unit cost.

(d) The error incurred is even smaller if w and $j(1+j^*)$ have smaller values than the ones assumed in generating the table. However, in the event that H and the responsiveness of I_W to changes in costs are high, an attempt to estimate h may well be justified.

CASE 4: A CHANGE IN V

As assumed in Table 1, a change in the volume V triggered by external forces has no effect on the unit price p, but it does affect the unit cost c. If we denote by c_o^* the direct change in c caused by the change in V, we have from (5)

$$c_o^* = -(1 - f_1)\frac{V^*}{1 + V^*}. \tag{20}$$

If we adopt the more general assumption (12) for the way in which the working

capital is affected, then the direct effect of V^* and the secondary effects are shown in the network diagram in Fig. 3, where the direct change c_o^* causes a change I_W^* which affects the bank loans and overdrafts B and then J and c, and so on.

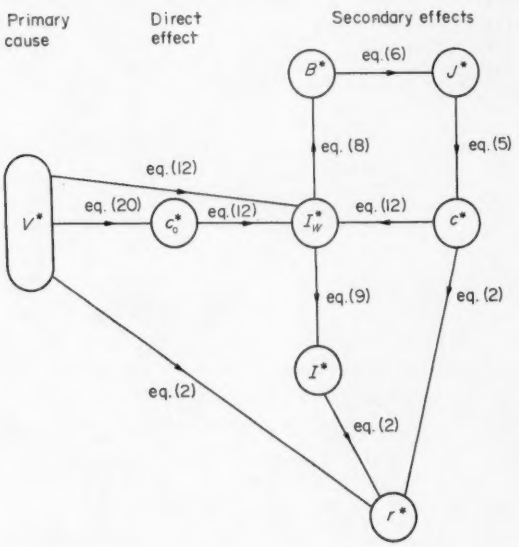


FIG. 3. *The effect of* V*—*case* 4 (*Table* 1) *for assumption* (12).

The iterative process culminates in a final value for c^* and hence for I_W^* and I^*, from which we can then determine (taking account of the primary cause V^* as well) the final effect r^*.

Thus, by (12) we find that

$$I^* = (g + h)V^* + hc^*(1 + V^*) = \frac{mV^*}{jH} \tag{21}$$

where

$$m = \frac{t + f_1 u}{1 - u}$$

$$t = jgH = jgI_W/(cV)$$

$u = jhH$ (as in (18) except that here $j^* = 0$)

Vol. 5
1977

so that the value of m can be determined.[4] From (8) and (6) we have

$$J^* = B^* = \frac{g+h}{\lambda} V^* + \frac{h}{\lambda} c^*(1 + V^*). \tag{22}$$

Using (5) we can now carry out the first iteration of the secondary effects and find the new value of c^* as

$$c^* = c_o^* + \frac{V^*}{1+V^*}(t + f_1 u). \tag{23}$$

We can proceed with the iterative process, or alternatively substitute $s^* = F^* = j^* = 0$ into (5) and then use (22) to solve for c^*, and in either case obtain the following result:

$$c^* = c_o^* + \frac{mV^*}{1+V^*} = c_o^*\left(1 - \frac{m}{1-f_1}\right). \tag{24}$$

Thus, (20) and (23) are approximations for (24). The solution for c^* for the alternative assumptions (10) or (11) can be obtained as a special case from (24): for (10) we substitute $h = 0$ (hence $u = 0$) whereas for (11) we substitute $g = 0$ (hence $t = 0$). If the secondary effects of a change in volume on the unit cost are ignored, i.e. if we assume $g = h = 0$, then $c^* = c_o^*$ as given in (20).

The limiting value of I^* can now be obtained from (9) as $I^* = wI_W^*$ so that

$$r^* = \frac{1}{1 + wI_W^*}\left[-\frac{1-a}{a}(1 + V^*)\, c^* + V^* - wI_W^*\right] \tag{25}$$

where c^* is given by (24) or by the approximation (23) and I_W^* by (21). On the other hand, if it is assumed that $g = h = 0$ (i.e. $c^* = c_o^*$ and $I_W^* = 0$) the expression in (25) is reduced to

$$r_o^* = \frac{V^*}{a}[1 - f_1(1-a)] \tag{26}$$

which gives the direct effect on r of a change in V.

The evaluation of (25) is summarised in Table 3. The numerical results indicate that:

(a) The sensitivity of r^* to g and even more so to h is relatively small, particularly for low values of g or h, so that an error in estimating them does not have a significant effect on r^*.

[4] It is not difficult to show that the ratio m can be expressed as

$$m = jH\,\frac{I_w^*}{V^*} = \frac{j\delta B}{c\delta V}$$

where the numerator $j\delta B$ is the incremental interest charge and the denominator $c\delta V$ is approximately the incremental cost, so that for small changes m becomes the marginal interest charge with respect to a unit increase in costs. Furthermore, it can be shown that as in this case $B^* = J^*$ it follows that

$$m = f_3\frac{J^*}{V^*}.$$

(b) If H and the anticipated responsiveness of I_W to changes in either costs or turnover are small, the effect on r of the secondary changes in c can be ignored and the simple formula (26) used instead of (25). It should be emphasised, however, that any conclusions drawn as a result of (25) or (26) are subject to the validity of the assumption that a change in the level of activity has no effect on the unit variable cost. In the event that the unit variable cost is expected to change, the analysis has to be modified accordingly.

TABLE 3. VALUES OF r^* (IN %) BY EQUATION (25) FOR $\alpha = 0{\cdot}10$, $w = 0{\cdot}5$, $f_1 = 0{\cdot}5$, $V^* = 0{\cdot}01$ AND $j = 0{\cdot}15$

(i) $h = u = 0$

H \ g	0	0·2	0·4	0·6	0·8	1·0
0·2	5·50	5·35	5·19	5·04	4·88	4·73
0·4	5·50	5·29	5·08	4·87	4·67	4·46
0·6	5·50	5·24	4·98	4·71	4·45	4·19
0·8	5·50	5·18	4·87	4·55	4·23	3·92
1·0	5·50	5·13	4·76	4·39	4·02	3·65

(ii) $g = t = 0$

H \ h	0	0·2	0·4	0·6	0·8	1·0
0·2	5·50	5·42	5·34	5·26	5·18	5·10
0·4	5·50	5·40	5·29	5·18	5·06	4·94
0·6	5·50	5·37	5·23	5·08	4·93	4·77
0·8	5·50	5·34	5·17	4·99	4·80	4·59
1·0	5·50	5·31	5·11	4·89	4·66	4·41

Vol. 5
1977

CASE 5: A CHANGE IN p

Over the short term, a small change in the unit price changes V by

$$V^* = -\epsilon p^* \tag{27}$$

where ϵ is the absolute value of the price elasticity of demand[5] at price p. As a result, the relative change in unit cost is as given in (20), namely

$$c_o^* = -(1-f_1)\frac{V^*}{1+V^*} = (1-f_1)\frac{\epsilon p^*}{1-\epsilon p^*} \tag{28}$$

an expression derived and discussed in [4].

The sequential effects of a change in unit price are shown in Fig. 4: p^* causes V^* which results in c_o^*; thereon the pattern is similar to the network in Fig. 3, except that here p^* has a direct effect on r^*. The relative change in working capital is by (12)

[5] If the change in p is relatively large, ϵ in (27) may be replaced by the estimate of the arc elasticity between p and $p+\delta p$.

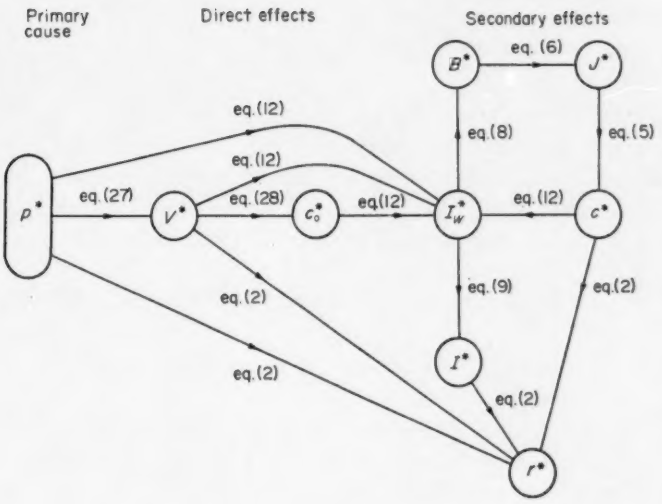


FIG. 4. *The effects of* p*—*case* 5.

$$I_w^* = (g+h)V^* + (1+V^*)(gp^* + hc^*). \tag{29}$$

It can further be shown that

$$I_W^* = \frac{mV^*}{1-u}$$

which is as derived for the previous case in (21), except that now[6]

$$m = \frac{t(1+p^*-1/\epsilon)+f_1 u}{1-u}$$

and

$$J^* = B^* = \frac{1}{\lambda} I_w^*.$$

Substituting J^* into (5), using (27) and solving for c^* we get

$$c^* = c_o^* \left[1 - \frac{t(1+p^*-1/\epsilon)+f_1 u}{(1-f_1)(1-u)}\right] \tag{30}$$

where c_o^* is given by (28). The result in (30) is similar to that in (24), which was obtained for Case 4, except that the first term in the numerator now reflects the direct effect that p^* has on the working capital and thereby on the unit cost.

[6] As in the previous case, it can be shown that

$$m = \frac{j\delta B}{c\delta V} = f_3 \frac{J^*}{V^*}.$$

It follows that

$$c^* = c_o^* \left(1 - \frac{m}{1 - f_1}\right)$$

which is the same as (24) with the appropriate definition of m. If p^* is relatively small it may be ignored in the numerator of (30), and a cruder approximation is obtained with $c^* = c_o^*$, which applies when $g = h = 0$. As in (24), the special case of assumption (10) is derived when $h = u = 0$ and for assumption (11) when $g = t = 0$ in which case the result is the same as that found by (24) for $g = t = 0$.

The change in the rate of return is now obtained from (2), after substituting $I^* = wI_W^*$

$$r^* = \frac{1}{1 + w\,I_w^*}\left\{\frac{1 - \epsilon p^*}{\alpha}\,[p^* - (1 - \alpha)c^*] - \epsilon p^* - wI_w^*\right\} \tag{31}$$

where I_W^* is given by (29) and c^* by (30). The special case $g = h = 0$ reduces (31) to the following expression [4], which may in appropriate cases be regarded as an adequate approximation

$$r^* = \frac{p^*}{\alpha}\,[1 - \epsilon\,(1 - f_1 + \alpha f_1 + p^*)]. \tag{32}$$

Vol. 5
1977

Numerical values for (31) are given in Table 4. As in the cases examined earlier, the results in the table clearly show that the secondary changes in I_W which are triggered by a primary cause of change in p have a small effect on r. The conclusion is again important in that no matter what assumption is made regarding the functional relationship between the working capital requirements and the general level of the firm's activity, the sensitivity of r to such an assumption is relatively small. Consequently, sensitivity analyses for planning purposes (for details see [3]) can often be performed in terms of quite simple models.

TABLE 4. VALUES OF r^* (IN %) BY EQUATION (31) FOR $\alpha = 0{\cdot}10$, $w = 0{\cdot}5$, $f_1 = 0{\cdot}5$, $p^* = 0{\cdot}01$, $j = 0{\cdot}15$, $H = 0{\cdot}4$

(i) $h = 0$

ϵ \ g	0	0·2	0·4	0·6	0·8	1·0
0	10·00	9·79	9·58	9·38	9·17	8·96
0·5	7·20	7·10	6·99	6·89	6·79	6·69
1	4·40	4·40	4·40	4·41	4·41	4·41
2	−1·20	−0·99	−0·78	−0·56	−0·35	−0·14
5	−18·00	−17·16	−16·32	−15·47	−14·63	−13·79

(ii) $g = 0$

ϵ \ h	0	0·2	0·4	0·4	0·8	1·0
0	10·00	10·00	10·00	10·00	10·00	10·00
0·5	7·20	7·25	7·31	7·36	7·42	7·48
1	4·40	4·51	4·61	4·72	4·84	4·95
2	−1·20	−0·99	−0·77	−0·55	−0·33	−0·09
5	−18·00	−17·47	−16·93	−16·38	−15·82	−15·23

In applying (31) or (32) one may be faced with lack of information as to the precise value of ϵ that should be adopted. However, it may often be sufficient for decision making purposes simply to compare the 'critical' price elasticity of demand

$$\epsilon_o = \frac{1}{1 + p^* - f_1(1-a)} \approx \frac{1}{1 - f_1(1-a)} \tag{33}$$

obtained by setting $r^* = 0$ in (32) and solving for ϵ, with what one would expect as an actual price elasticity of demand. Thus, (32) may be rewritten as

$$r^* = \frac{p^*}{a}\left(1 - \frac{\epsilon}{\epsilon_o}\right) \tag{34}$$

When ϵ/ϵ_o is significantly smaller or larger than 1, precise knowledge of ϵ is not necessary, if all that management is concerned with is to know whether a change in unit price is likely to increase or decrease r. If, on the other hand, there is reason to believe that $\epsilon \approx \epsilon_o$, it is evident from (34) that only if profit margins are low would an attempt to estimate ϵ accurately be justified, otherwise the sensitivity of r to ϵ is small. Table 5 gives values of ϵ_o for $a = 0{\cdot}20, 0{\cdot}10, 0{\cdot}05, 0{\cdot}02$, and $1-f_1 = 0, 0{\cdot}2, 0{\cdot}4, 0{\cdot}6, 0{\cdot}8$ and 1.

TABLE 5. THE CRITICAL PRICE ELASTICITY OF DEMAND

a \ $1-f_1$	0	0·2	0·4	0·6	0·8	1
0·20	5·0	2·8	1·9	1·5	1·2	1·0
0·10	10·0	3·6	2·2	1·6	1·2	1·0
0·05	20·0	4·2	2·3	1·6	1·2	1·0
0·02	50·0	4·6	2·4	1·6	1·2	1·0

CONCLUSION

The results for the five cases listed in Table 1 are summarised in Table 6, where a reference to the corresponding equations in the text is given.

TABLE 6. RESULTS FOR c^* AND r^*

Case (see Table 1)	Primary cause of change	Formulae for c^*		Formulae for r^*	
		Final result	Approximation	Final result	Approximation
1, 2, 3	s^* and/or F^* and/or j^*	(16)	(13)	(19)	(14)
4	V^*	(24)	(20)	(25)	(26)
5	p^*	(30)	(28)	(31)	(32)

While the analysis in this paper is directed to an investigation of the effect of each primary cause of change, it is possible to use the same methodology for an

examination of several simultaneous primary changes. This was partly done in the paper in combining cases 1, 2 and 3 by considering s^* and/or F^* and/or j^*, and the analysis can easily be extended to a more detailed decomposition of the unit cost structure, if that is thought to be desirable. As imilar approach may be employed for the general case, which encompasses all the cases described in Table 1, in order to provide a general solution to the problem of establishing the total effect of primary changes on the unit cost, on profit and on the return on investment.

REFERENCES

1. EILON S (1975) Changes in profitability components. *Omega* **3** (3), 353–354.
2. EILON S (1976) Sensitivity of the rate of return. *Omega* **4** (3), 350–352.
3. EILON S, GOLD B and SOESAN J (1976) *Applied Productivity Analysis for Industry*. Pergamon Press, Oxford.
4. EILON S and PHOCAS-COSMETATOS G (1976) Analysis of profitability components. Proc. Euro II, Second European Congress on Operations Research, pp. 143–147.
5. GOLD B (1971) *Explorations in Managerial Economics*. Macmillan, London.
6. SAMUELS JM and WILKES FM (1975) *Management of Company Finance*. Nelson, London.

ADDRESS FOR CORRESPONDENCE: *Professor Samuel Eilon, Department of Management Science, Imperial College, Exhibition Road, London SW7 2BX, UK.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 6, 1977. Pergamon Press. Printed in Great Britain

The Management of Timber and Related Industries in Sweden

HG JONES

The University of Strathclyde, Scotland

(*Received February 1977: in revised form March 1977*

Timber is a major raw material in Sweden and legislation mostly concerned with the continued viability of the industry has been enacted. Interesting organisation structures to deal with varying circumstances have evolved.

INTRODUCTION

ALMOST 60% of the land area of Sweden is devoted to forestry, much of which constitutes a major recreational asset with a mixed sylviculture, containing silver birch and some oak as well as pine and spruce. Right of access to the forest lands is part of the common law of Sweden.

Of course, the forests are more than an amenity. About one quarter of the country's exports by value are based on forestry products. With characteristic logic, the Swedes have recognised that it pays to export products with a high added value. With equally characteristic realism, the primary producers have determined that they, too, should benefit from the high prices obtained for the end product of their industry; the result is widespread integration, almost from seedling to market.

The days of a small man doing everything himself are gone. Modern pulp and paper mills require considerable capital investment, and Sweden is fourth in the production of the world's wood pulp. The old sawmills cannot compete in price or quality with their modern counterparts. Sophisticated furniture and prefabricated timber houses of quality design both require effective marketing to derive maximum profitability. But before discussing the management methods by which Sweden has achieved these ends, government action in the economic field merits attention.

LEGISLATION

Successive governments throughout the twentieth century have been concerned to preserve the long term viability of the forestry industry in Sweden. Thus in early years it became a legal requirement to plant new woodlands after a felling. A Forestry Board was established in 1903 to monitor performance through 24 county boards which now have to grant permission for an owner to fell a stand of timber. Young forests have been protected since the 1920's by a law which states that the only felling therein can be by selective thinning. In order to maintain a high yield, the Forestry Act of 1948 went still further by defining the size of trees which can be felled.

Nineteen seventy-four was a year of peak demand in which marginally more timber was felled than replanted. The Socialist government was then considering legislation to ensure that a new tree was planted for every one felled. The forestry owners argue that such legislation is unnecessary because the forces of economics will dictate the same policy and in any case, the slackening of demand due to world recession in 1975 and 1976 redressed the planting deficiency of 1974. It would appear that the capacity of the existing forests to meet demand is nearing saturation level at a total annual volume of about 7·8 M cubic metres worth £1000 M in final product; many finishers feel they are unlikely to be able to increase output along conventional lines and so are experimenting with methods of improving the yield by utilising brushwood and tree stumps or roots for example. A voluntary agreement has been made within the industry to limit industrial expansion up to 1978.

Vol. 5
1977

To encourage further planting, a law passed in 1965 enabled finishers to acquire land for forestation but so far relatively little use has been made of the facility, presumably because private capital has been needed to finance the new very large pulp mills. The government comes within the category of an owner of finishing plants so under this Act, it, too, can purchase land for forestry, a feature which the small owners view with distaste as being a potential "rape of the industry" to use their emotive words.

The Swedes are intensely safety conscious. Safety, supported by legal sanctions, has been a main reason for the rapid development of mechanical equipment for use in the forest although this has paid off handsomely by increasing the output per man day some 20-fold over the last two decades.

Safety was also the reason for the government banning the use of DDT for the protection of seedlings. The larva of the snout beetle breeds in the bark of the stumps of newly felled trees: when the beetle emerges, it attacks any tree seedlings in its vicinity. If the seedlings are protected with DDT, the ground can be replanted within 6 months of the felling but without DDT, 3 or 4 years need to elapse for the bark to rot and for the beetles to perish. Thus the loss of 2½ years of timber growing is set against the risk of infection of workers in the forest nurseries. In contrast, at the other end of the growing cycle, nitrogenous fertilis-

ers, applied from aeroplanes, in lumps the size of sugar cubes some 20 years before the end of the normal growing time (of 60–80 years in the South, compared with 120–140 years in the North) can reduce the growing time by 10 years: such fertilising is becoming standard practice in the search for efficiency. In mid 1975, the government repealed a ban on the use of chemical spraying from the air for controlling tree types but to date little seems to have happened in this direction.

Protection of the environment has become a major issue in Sweden. Effluent from the pulp and sulphite mills has been a particular target in the 1970's and has been the subject of legislation.

DIFFERENCES IN THE NORTH AND SOUTH

The growing time is not the only point of difference between forestry in the North and South of Sweden. In the North, the owners are mostly large companies (25% of the total forestry area in Sweden is worked by large companies) or the State and municipalities (another 25%) whereas in the South small private owners predominate.

Different practices, too, have evolved. In the North, clear felling is employed, i.e. the land is stripped of timber with heavy mechanical equipment—a mobile integrated machine for this purpose can weigh 25 tons. In due course the land is replanted with graded seedlings to improve the timber stock: as a result of research, some nurseries can produce in 4 months seedlings of a size that previously took 3–4 years. In the South, the private owners often work one-man farms and elect to do their own felling in the winter time with hand operated power saws; as often as not, the stock is improved by selective thinning. Full time lumbermen are also employed in the South but the equipment is generally lighter although the trees are lifted, debranched and cut to logs by mobile equipment operated by only one man. Sometimes the teams in the South will be full time workers together with owner farmers working part time.

It is partly because of the more severe weather conditions in the North that the process is more completely mechanised to include in addition to actual felling, the automatic grading and loading of logs for transport to the mills. Even so, because of the weather it is more uncomfortable and more difficult for the workers to achieve steady bonus performance. As a result, the pressure came from the workers in the North for the abandonment of bonus schemes in favour of fixed monthly wages, and in fact led to unofficial strikes in 1974. The lumbermen in the South refused to join the strike; the younger men in particular preferred the higher take home pay they could earn from a bonus scheme. The union negotiated monthly pay for North and South. So far there are rumours of loss of productivity in the North of up to 20% but in the South figures of a half this magnitude—or even less—are quoted because workers now need to stay on the

job all day instead of going to own activities after the bonus rate has been achieved.

Eighty-six per cent of the equipment used in the forest is less than 10 years old. It is purpose built for the job and is operated with sophisticated controls located in air conditioned cabs of ergonomic design. Moreover, the performance and profitability of a particular type of machine are not considered in isolation; each machine is seen as part of a total logging system in which terrain, planning and training of personnel all influence the situation. Displaying their usual commercial acumen, the Swedes have developed a healthy export trade for the sale of this equipment.

A NORTHERN BRUK

The private companies in the North are usually based on that uniquely Swedish organisation known as the bruk. Throughout the seventeenth and much of the subsequent two centuries, the burning of trees for charcoal to be used in the smelting of iron ore mined locally was carried out in many small communities with the landowner being very much responsible for everything and everybody within the community. As technical progress advanced, the manufacture of iron and steel moved into progressively larger and fewer units, while the timber was fed first into sawmills (which also had a similar progression in size and number) and later into pulp mills also. In spite of the changed structure, the bruk in most cases was able to maintain its twin interests of steel and forestry. Thus a company such as Uddeholm, for instance, is technically very advanced in both fields and in addition has forty hydroelectric installations and a chemical division.

The diversification is believed to be beneficial in evening trade cycles and also in the timing of the deployment of capital. In 1975, which was not a good year for exports, the total operating income for Uddeholm was made up in the ratios of chemicals 1, power 2, forestry products 2 and steel 3. For the peak year of 1974, the income from chemicals and power was approximately the same but the contribution from forestry products on this scale was 8 and from steel 6.

With a total of 13,500 workers, Uddeholm dominates the economy of the fairly remote county of Wermland in which it is situated, but 60% of its sales are in export markets. The fortunes of this bruk, started in 1668, were steered by generations of the same family before it became a public company in 1870 in order to raise capital for a new and bigger blast furnace. The latest development, made under a new managing director Gunnar Wessman, is towards decentralisation in order to cope with the rapidly changing and diverse demands of the several special steel producing units; at the same time as decentralisation for operating purposes is introduced, a strong central control of finance is maintained.

The Uddeholm bruk still manages more than 1·1 million acres of mixed

Vol. 5
1977

forestry and agricultural land (80% under forestry); forestry involves the same order of capital investment as the steel division being nearly £20 M in 1975 (although the 1975 forestry division investment was exceptionally large because of the construction of a new board machine). The investment can be seen in perspective over a longer period such as the last five years for which the total investment in fixed assets has averaged 8·6% of invoiced sales (which were £260 M in 1975); the source of the extra capital has been long term loans and a rights issue in 1976 which raised an amount equal to a half of the 1975 investment.

From the monolithic nature of the organisation, it is possible to phase tree felling to the demands of the mills without the necessity for intermediate stock-piles in the forests—a computer system employed in the planning contains data for every stand of timber owned by Uddeholm. Associated with this modern management technique is one of the oldest transport systems in the world—about 25% of the timber, amounting to 10 million logs a year, is floated down the river Klarälven to the mills. Some of this timber is imported from Norway to meet the increasing demands of the mills. In fact Uddeholm can now only supply from their own forests about 50% of the mills' requirements.

Women make up 40% of the total work force in Sweden. Forestry and steel making are both traditionally male preserves so that in an area dominated by these industries, job opportunities for women are considerably less than the average. This is a problem exercising Uddeholm and although women are encouraged as far as possible to enter these industries, the job potential seems to be less than 20% of the labour force. The difficulties are particularly acute when recruiting trained technologists because the lack of job opportunities for wives becomes an important factor.

A SOUTHERN CO-OPERATIVE

Faced by competitors possessing resources on the scale described, the 44,000 small timber owners, with a combined forest area of 5 million acres located south of an irregular line drawn from Gothenburg to Norrköping, needed a mechanism for economic defence; the problems of scale have been tackled through a co-operative organisation owned by the farmers but nowadays run by specialists. The origin of this co-operative dates back to only 1933, and is largely due to the inspiration and energy of a remarkable man named Gösta Edström. It is a favourite saying in Sweden that initially Edström had to borrow the money to pay for the stamps. Certainly he travelled the countryside on a bicycle to persuade the independently minded farmers to sell timber to him at an agreed price for resale to the sawmills. By 1944, the co-operative (known as SSSF) was able to buy its first small sawmill. A paper mill at Strömsnäs was acquired in 1955 and from then on the enterprise flourished so that when Edström retired as chairman

and managing director in 1973, annual invoiced sales were nearly £150 M and the number of employees was 6800.

Capital for the co-operative is raised from each owner at the rate of 100 Skr. per hectare (approx. £6 per acre) of his land under forest but only about half of this exists as fully paid up capital. When a farmer fells a stack of timber (or also when it is felled for him by SSSF), the timber is stacked near a track and an independent inspector is notified. The inspector verifies that the amount of timber is in accordance with the approval granted by the Forestry Board and attaches a metal label to the stack showing the amount and the map co-ordinates of the stack; this information is fed into a central computer system of SSSF.

Because of the structure of the organisation, it is not possible to phase the felling in accordance with the demands of the mills; hence a stack of timber may lie in the forest for several months before being transported to the mills. However, the farmer receives a first payment within 30 days of the inspection but 4% of the value of the timber is allocated to his capital credit—a sore point is that this capital is taxed as income. The computer system co-ordinates the picking up of the timber, a second check being made on volume of usable timber from each stack, and at the end of the year a second payment is made to the farmer depending on the profitability of the whole enterprise and the value of the timber delivered from his stack. The computer is at the centre of these activities and again exhibits a neat application of a modern management technique in one of the oldest occupations in the world.

Vol. 5
1977

Unlike the Uddeholm bruk with its diversified interests, SSSF has achieved a substantial turnover by concentrating on timber based products mainly in order to provide profitable outlets for timber supplied by the members. The mechanical division now comprises 13 sawmills, 3 chipboard mills and a factory producing 2000 timber framed houses a year. Some of the equipment in the sawmills is 20 years or so old and still works adequately but when the quality of the product as it reaches the customer is involved, investment has drawn upon modern technology in the shape of controlled drying kilns, and of machines for stress grading structural timber, and the sorting, banding and labelling for dispatch in order to ensure customer satisfaction.

Investment in SSSF has averaged 10·9% of turnover during the last 5 years. The chemical pulp division has been a large consumer of capital for the provision of three major pulp mills each located on the coast in order to ease the disposal of effluent in an acceptable manner. Although capital intensive, the two newer pulp mills have a turnover of nearly £100,000 per employee; in comparison the sawmills have a turnover of £36,000 per employee, a figure which is also gratifyingly high. With the restriction on felling, expansion in pulp mills in future is likely to be limited but there may be development in paper manufacture for which SSSF already has one mill, a unit producing paper sacks and another works manufacturing cardboard and cardboard boxes.

In the early days, SSSF had difficulty in raising loans but with established

assets showing a pre tax return on working capital of 15%, the current position is very different, with the Investment Bank prominent among the lenders. Outstanding loans and mortgages amount to 16% of the total assets.

Part of the strength of SSSF is derived from the technical assistance given to members for the management of their forests as well as in the provision of seedlings of improved quality. Each year, a considerable effort is made to involve members in area discussions of factors influencing the performance of the industry. As many as 5000 members took part in 1974 in a 'Top Marks' competition for knowledge about forestry.

MANAGEMENT PHILOSOPHY

Many companies in Sweden are diversified to a surprising extent. Diversification has arisen partly because of the need to absorb workers displaced by the process of modernisation and partly because of a keen awareness of the market potential of spin-off technology; in addition there has been the normal desire to spread risks and to be at the profitable end of the business. The majority of companies are large enough to make decentralisation into operating divisions a sensible as well as a fashionable procedure. In most cases, the head office centre maintains a strong control through the planning and finance functions, both of which have been developed to a sophisticated level, usually with only a few people who display considerable competence.

Apart from a difference in time scale for long term planning (when forestry needs to think in terms of as many decades as other industries do of years), the two companies in the timber industry already discussed although very different from each other in structure, reveal the typical characteristics of Swedish management. The long term viability of the dependent industries is a main concern. The fact that felling is currently in approximate balance with new growth conceals a reserve derived from a growing timber stock of about 2400 M cubic metres [1] of maturing timber, mostly in the South, created over the years 1920 to 1972 when growth exceeded felling. Nonetheless, this reserve could easily be dissipated by unwise policies. Forward planning concentrates on methods of improving growth in the forests, and on methods of improving the yield (presently about 84%) by exploiting new methods of utilising bark, branches and even stumps and roots of the trees. The balance between the allocation of timber to saw or pulp mills is kept under careful review for maximum profitability: any major increase in labour costs could influence the balance and might also lead to fundamental changes in the production methods in the forest. Major increase in existing pulp capacity is unlikely so the companies are examining the effect of extending manufacture into the finishing processes (e.g. paper, where the profit margins are less than with pulp), when they could be entering into competition with existing customers. Enough has been said to

create a verbal model of a fascinating problem confronting the strategic planners.

The short term company plan, including the one year budget (usually expressed in terms of men and materials as well as money), consists of the aggregation of plans prepared by the divisions for discussion and approval by the centre. It is widely recognised in Sweden that the effectiveness of the budget as a control instrument is largely determined by the skill and care with which it is prepared; there is concern that the mechanics of budget preparation should be better understood. It is also accepted that objectives will differ at different levels within a company. All these general points are apparent in the cellulose division of SSSF where the controller, Arne Carlsson, employs an information flowchart as a check list during the consultation that takes place over several months from autumn to spring during the preparation of the budget. The chart contains 'boxes' for the objectives of the various functions (sales, production, etc.) at different levels within the division's hierarchy; constraints too will vary throughout the division, some being internal and therefore to some extent adjustable while other constraints such as the state of the market, are external and uncontrollable by the company but nevertheless react with the division. The factors imposing limitations on change are regarded as being in special need of clear identification. The next stage is to integrate the objectives into a compatible and acceptable pattern. Then the operating data must be fed back in such a form that each section can readily assess its performance against its objective; the intention is that corrective action would then be self motivated. The administration computer system (of which only a small part was operational in the autumn of 1976) is designed to give the information flow. The main problem lies in the difficulty of being explicit on both objectives and how the routine data should be analysed. The attention given to this aspect of the budget is interesting evidence of the professionalism of Swedish management.

Vol. 5
1977

Industrial democracy thrives in Sweden. Two representatives of the workers sit on the boards of most companies. Consultation through the Works Councils is widespread and effective. In rural areas, for example in the forest and sawmills, consultation is often also informal (and sometimes paternalistic), management conceiving the supply of correct information to the workers as being an important part of the job. Company economics are explained to many workers through the balance sheet.

Forestry throughout Sweden is supported by the activities of the Technical University for Forestry, presently located just outside Stockholm. On the management side, one aspect of this institution's work is the introduction of economic thinking into traditional planning by developing a logical system to establish criteria by which certain options can be eliminated from further consideration. The linear programme for a mixed farming enterprise is of interest because of the detail introduced in order to give realism to the model [2]. The capital yield at the end of the planning period (of 25 or 30 years) is maximised subject to the constraint that the enterprise must yield sufficient income for

the farmer to live throughout the period. The input includes the amount of labour the farmer and his family can devote to the various activities in different years, the cost of hired labour and provision for the replacement of equipment, as well as the size and age distribution of the opening stocks of standing timber. Even with this amount of detail, the University accept that the model cannot be used for the actual planning of a particular farm but it can give useful guidelines which can be of greater value if a group of farms are under consideration.

Forestry represents a major item in the Swedish economy. The distinctive approaches to success discussed here, and the general climate of progressive thinking augur well for the future.

REFERENCES

1. KILANDER K (1972) Swedish forest potentials. Värö Conference Paper.
2. VON MALMBORG G (1969) Economic planning of the farm forest operating unit. *Studia Forestaliasuecica* No 71, 1–56.

ADDRESS FOR CORRESPONDENCE: *HG Jones Esq, University of Strathclyde, Dept of Operational Research, Livingstone Tower, 26, Richmond Street, Glasgow G1 1XH, UK.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5. No. 6, 1977. Pergamon Press. Printed in Great Britain

Market Representation in the Motor Industry

JDW MORECROFT
Massachusetts Institute of Technology

CDT WATSON-GANDY
Imperial College

(*Received June* 1977)

A model is described to determine the best location for motor car dealers and the area they should serve. The method described uses a spatial interaction model and measures sales potential. The model has possible application with other consumer durables or in locating salesmen.

INTRODUCTION

A SIGNIFICANT part in the marketing strategy of an automobile manufacturer is the location of dealers and the demarcation of the areas these dealers will serve. The process of assessing the network of dealers is a continual process as new towns grow and as existing towns and their inhabitants change in character and purchasing behaviour.

Any study of dealer marketing areas will normally fall into three parts. First, a field study is made of a region which may include several possible dealer areas. This involves driving around in the region noting the locations of different qualities of housing (council, high class residential, etc.), the road network, details of the shopping centres, the location of competing dealers, and so on. The second part of the survey is an assessment of the field work, together with population and income statistics, and planning information if this is available. This work culminates in a report making recommendations for changes in the dealerships and the demarcation of the areas. The final part of the study is the consideration by higher management of the report together with other qualitative and quantitative factors. These factors include not only the costs of the changes

but also the relationships with existing dealers, many of whom have invested substantial amounts of money in their business. A further consideration in the case of proposed new dealerships is whether there is sufficient business to support a dealer. It is obvious that complete saturation of an area will initially sell more cars, but if there are insufficient sales to the individual dealers, many will go out of business and therefore be unable to supply parts, servicing and the after sales service necessary to keep customer loyalty.

This three stage approach to the problem has many strengths and weaknesses. The principal strength of the method is its flexibility in that it allows for the introduction of many variables from a wide range of sources. Further, the decision maker is only mildly constrained by the recommendations. The principal weakness of the system is that the variables need not be integrated into the report in a consistent way. The objective of the model presented here is, therefore, to make use of information on shopping centres and socio-economic data in a systematic manner and to provide an underlying structure on which to base the field survey. To do this the model will be used to indicate possible changes to the boundaries of dealer areas, the most suitable area to site a new dealership, and possible localities for secondary representation, thus relieving the staff for other work, for example, more extensive field work.

The model will be described below and illustrated by the results of an example study.

Vol. 5
1977

THE MODEL

The model has two component parts. These are the definition of the catchment area and the location of the dealership at the point of maximum sales potential.

Catchment area

The law of retail gravitation or Reilly's Law [12] states that the attractiveness (A_a) of a centre a to a potential customer is directly proportional to a measure of the size of that centre (S_a) (Reilly suggested population), and inversely proportional to a function of the distance (D_a) separating the customer and the centre. That is:

$$A_a = k \frac{S_a}{D_a^{\lambda}}, \tag{1}$$

where k = constant,

λ = an exponent (Reilly suggested $\lambda = 2$).

From equation (1) the point of equilibrium in terms of attractiveness between two centres can be found. Carroll [4] extended this work to show that the area of influence around a centre was circular (called the circle of Apollonius) but

with the centre of the circle surrounding a smaller centre displaced along the line joining the two centres. Carroll gave formulae (based on Reilly's Law) to calculate the radius and the displacement. This work was further extended by Dixon [6] who examined the interaction of more than two cities and gave more general formulae for the radius and displacement of the circles of Apollonius. These are:

$$\text{Radius } (R) = \frac{mD_{ab}}{m^2-1} . \tag{2}$$

$$\text{Displacement } (E) = \frac{m^2 D_{ab}}{m^2 - 1} \tag{3}$$

where D_{ab} = distance between centres a and b

$$m = \left(\frac{S_a}{S_b}\right)^{1/\lambda}$$

Much argument has centred on the validity of Reilly's Law (for example, see Scott [14] p. 169). However the model is widely used (see Kilsby *et al.* [10], or the Haydock study [11]), although a range of parameter values are selected and different definitions of centre size and separation are used. For example, in selecting the parameter λ, Reilly found values ranging between 0·0 and 12·5. In a particular study of a specified market, λ may be varied in Dixon's equations (2) and (3) to find the most appropriate value for the situation in hand. In the definition of centre size, population, retail turnover and retail floorspace are among many proxies which have been used (see Kilsby *et al.* [10]). In the example study mentioned later an index of size was constructed based on a weighted average of the total number of shops in the main shopping centre. Up to date and readily available information [16] determined the choice of this proxy. Finally, the definition of centre separation has been alternately stated in terms of economic distance (travel time) and real distance (miles). The former is probably a more accurate measure [13], but real distance was used in the example study due to availability of consistent data.

Dixon gives examples in which he examines the interaction of many centres but he does not give any indication of his method of allocation when circles of competing centres overlap. Several methods are suggested by Acar [1]. However, an iterative method is proposed which uses discrete population areas (wards and civil parishes are most suitable as these are the smallest units for which economic data is available). The method works as follows.

The centres are first arranged in descending order of size. The largest (dominant) centre is allocated all wards and civil parishes. The program then works through the hierarchy of dominated centres and reallocates the discrete areas on the basis of the interaction between dominant–dominated pairs. When this procedure is finished the catchment area of the first centre is defined by those

areas it still holds, and this cannot be changed by any secondary interactions between the dominated centres. The procedure is then repeated with the second largest centre and so on until all centres have been examined.

Market potential

The market potential of a location within a catchment area may be defined as a measure of the size of the market relative to the position of the retailing outlet. The fundamental aim of a market potential model, therefore, is to locate retailing outlets such that the economic distance to all customers within the catchment area is minimised.

The concepts of market potential and spatial interaction were developed by Beckman [2] who separated market potential for a retailer into two components, a locational potential and a potential of population. Beckman's locational potential (L) is defined by:

$$L = \Sigma_j \frac{A_j}{D_j} \tag{4}$$

where A_j = the attractiveness of outlet j;

D_j = the economic distance of the population centre to outlet j.

This leads to a market share for retail outlet i, defined by:

$$M_i = \frac{A_i/D_i}{\Sigma_j A_j/D_j}. \tag{5}$$

Beckman defines the potential of population as:

$$\pi_i = \int \frac{P(r)}{D_{ir}} \, dr \tag{6}$$

where $P(r)$ = the population density at point r

D_{ir} = the economic distance from retail outlet outlet i to point r.

Taking the effects of market share into account the market potential for retail outlet i is given by:

$$\phi_i = \pi_i \left(\frac{A_i/D_i}{\Sigma_j A_j/D_j} \right). \tag{7}$$

Some care is required in adapting this expression to represent potential demand faced by a car dealer. In particular the market share which a dealer can expect is not solely a function of fractional locational potential. It is related to the composition, by make, of the stock of cars in the market, the product differentials between vehicles of competing manufacturers and many other variables which characterize competition in the automobile market. Indeed, these consumer

Vol. 5
1977

preferences can constrain the value of the market share of a particular manufacturer to $\pm 5\%$ of his national average.[1] In short, market share cannot be calculated from Beckman's locational potential formula, although, in the long run, locational potential is one of a number of important determinants (otherwise market representation studies would be futile). To overcome this problem a regional average of market share (s) of the franchising manufacturer is used to apportion population potential between the different competing dealers.

The modified equation for market potential at the point i is therefore:

$$\phi_i = \pi_i s.$$

One further modification can then be introduced to the equation above to account for the discrete population distributions used in the model and at the same time to introduce a general distance attenuating function. The market potential of location i is therefore given by:

$$\phi_i = s \sum_j \frac{\sigma_j}{f(d_{ij})} \quad , \tag{8}$$

where σ_j = the demand for cars in ward j—determined on the basis of economic data and market research surveys.

$f(d_{ij})$ = a function of the distance between dealer location i and ward j.

Equation (6) is remarkably similar to the potential at any point on a force field in electrostatic theory [3, 9]. This analogy has the advantage of providing a positive charge (a demand for cars) at points on the plane which can be satisfied by introducing a negative charge at the dealer location (equivalent to the supplying of cars). This model can then be used not only to determine the best location for a dealer, but also to examine the usefulness of secondary dealer representation. Thus if a dealer is located at centre a, and the possibility of locating a secondary dealer at centre b is to be considered, then the market potential for dealer a without secondary representation is given by:

$$\phi_{ab} = s \sum_{j \epsilon A \cup B} \frac{\sigma_j}{f(d_{aj})}$$

where A and B are the sets of population units allocated to centres a and b respectively.

The market potential with dealers at both a and b is:

$$\phi_{a+b} = \phi_a + \phi_b$$

$$= \left(\sum_{j \epsilon A} \frac{\sigma_j}{f(d_{aj})} + \sum_{j \epsilon B} \frac{\sigma_j}{f(d_{bj})} \right) s.$$

[1] Based on experience of market penetration data from the Society of Motor Manufacturers and Traders.

The additional potential gained by locating a dealer at b is given by $(\phi_{a+b} - \phi_{ab})$. If this value is sufficiently large to first support a secondary dealer, second justify the necessary investment and third, the loss of sales to the dealer at a is not too great, then consideration can be given to secondary representation.

The distance function

Hlavac and Little [8] suggest from a study carried out in Illinois on the urban automobile market that the distance function is exponential. This is supported by studies carried out for the example reported here, in which the sales of new cars per economic head of population were plotted against distance from the dealer. These studies gave an equation of the form:

$$\frac{1}{f(d_{ij})} = Ke^{-kd}ij$$

Where K and k are parameters; K will reflect the regional penetration of the dealer, and k varies with the topographical nature of the dealer area, i.e. urban or rural, and the proximity or otherwise of other dealers.

Dealer location

Equation (8) can be solved with respect to the x and y coordinates to give the best location for maximum sales potential using an iterative procedure similar to that described for depot location in [7]. However, in an urban context with the consequent planning restrictions, precise grid coordinates have limited value. It is, therefore, suggested that a simple grid search over the area of interest is adequate for the purpose. In this way the best area for a dealership and the total area beyond which a dealer would be unfairly handicapped can be identified. The potential at a current dealer site can also be found and compared with the best potential obtainable. Hence an assessment can be made of the advantages of moving the dealer to a potentially more profitable locality.

Vol. 5
1977

THE ALGORITHM

The algorithm can now be usefully summarized into the following steps.

1. Select the centres. These might initially be the locations of existing dealers.
2. Define the catchment areas of each centre using the allocation procedure described earlier.
3. Determine the locality of maximum sales potential.
4. Determine the increased sales from secondary representation.

From a study of the results and from field surveys an assessment can be made of the advantages and disadvantages of moving current dealers and of the possibilities for secondary or new dealerships. These changes are made to the

input data and the procedure repeated iteratively until no further improvement can be made.

This procedure has two main advantages in that the speed of the computer is used to assess quickly the quantitative results of changes, while the results can be adjusted in the light of qualitative knowledge and experience outside the machine.

AN EXAMPLE

The area selected for study was the Birmingham and surrounding conurbations. This area was selected for two reasons. First, it forms a large and interesting market of some complexity; second, it was possible to obtain additional information on this area both as data and for comparative purposes. This region is divided into some 540 wards and civil parishes, and population and socio-economic data for these small administrative areas were obtained from the Small Area Statistics of the 1971 10% Sample Census [15]. The census gives details of the number of economically active people in 17 different categories. Data on the potential sales of new cars was obtained by grouping the original 17 categories into the A, B and C classes along the lines laid down by the British Bureau of Television Advertising and then weighting them by suitable market survey figures.

Fifteen major centres in the area were selected initially and their index of attractiveness calculated, as described previously, from the number of shops and department stores [16].

Using the methods described earlier the catchment area of each centre was calculated. Several values for λ, the exponent in equations (2) and (3), were tried. However, a value of $\lambda = 3$ gave intuitively the most satisfactory division of the area (see also the Haydock study [11]).

These calculated areas showed a broad agreement with existing dealer areas, but with some differences which can be implemented to improve retail coverage. Some small anomalies were also found. These could be due to the shapes and characteristics of the original ward areas or alternatively this suggests that the correct value of the parameter λ has not been used and that a higher value should be tried. However, a closer survey of the area would reveal whether or not the correct decision had been made. It is interesting to note that the calculated areas were, perhaps not surprisingly, in much closer agreement with a recent shopping survey of the area [5] than the existing dealer areas.

Some sensitivity analyses were performed to study the robustness of the solution to changes in the attractiveness of different centres. These indicated that the calculated areas are relatively insensitive to small (up to 15%) changes

in attractiveness, but also illustrate the ease with which the effects of possible changes, e.g. a proposed new shopping centre, can be analysed.

The next stage in the algorithm is to calculate the sales potential at dealer points. This was done using the methods described earlier. Equation (8) in the form:

$$\phi_i = sK \Sigma_j \sigma_j e^{-kd}ij \tag{9}$$

was used for this purpose and also in the form:

$$\phi_i^* = sK \Sigma_j \sigma_j \tag{10}$$

ϕ_i^* from equation (10) gives the saturation level for area i and is used to determine the percentage coverage of the market area. The values for K and k were derived from the study described earlier into the distance attenuating function.

It was difficult to obtain figures with which to compare the sales potential as calculated by the model with actual sales. However, based on a very small sample, both the predicted sales and the percentage coverage seemed lower than expected. One reason for this is that there is a degree of overlap in sales areas obtained in practice and this is not represented in the model. Changing the value of k in equation (9) improved this situation. Closer inspection of the percentage coverage showed that although coverage was good in urban areas it was less so in rural areas. The model also suggested several points for secondary representation (some of which are now under consideration), implying that the minimum potential sales constraint on new outlets may be too severe. Relaxing this constraint would certainly improve coverage, but this needs careful consideration in the light of the company policy on the service to be provided to the consumer.

The final step in the algorithm, the grid search, showed that in all cases where dealers already existed, they were sited within approx. 1 km. of the indicated point of maximum potential. For areas where no dealer existed, the model indicated the preferred regions for closer examination by field survey.

Vol. 5
1977

CONCLUSIONS

A model has been described to determine the best locations to place dealers and to delineate the areas they should serve. The model is proposed to assist management in making decisions on dealer representation in a consistent manner nationally, particularly with the use of the data obtainable from [15] and [16].

The model is based on the automobile industry, but has possible application to the examination of representation in many other consumer durables. With modification it can also be applied to the delineation of salesman territories.

Further work on the distance attenuating function i.e. determining the values of K and k in equation (9), could be used to obtain an interesting insight into the market and marketing skills of individual dealers.

REFERENCES

1. ACAR W (1974) A method of allocating recreation supply to urban centres. *INFOR* **12**(3), 231–244.
2. BECKMAN MJ (1971) Market share distance and potential. *Regional Urban Ec.* **1**(1), 3–18.
3. *Electricity and Magnetism* (1965) Berkeley Physics Course, Vol. 2. McGraw-Hill, New York.
4. CARROLL JD (1955) Spatial interaction and the urban-metropolitan description. Papers and Proceedings of the Regional Science Association, 1, D.
5. *City of Birmingham Structure Plan* (1973) Report of Survey—Shopping.
6. DIXON OM (1972) Models of nodal regions, the retail structure of cities. Occasional paper No. 1, Urban Study Group, Institute of British Geographers.
7. EILON S, WATSON-GANDY CDT and CHRISTOFIDES N (1971) *Distribution Management: Mathematical Modelling and Practical Analysis*, Chap. 4. Griffin, London.
8. HLAVAC TE and LITTLE JDC (1966) A geographic model of an urban automobile market. Working Paper, Sloan School of Management, M.I.T., 180–66.
9. JACKSON JD (1967) *Classical Electrodynamics*. John Wiley, New York.
10. KILSBY DJE, TULIP JS and BRISTOW MR (1972) The attractiveness of shopping centres. Occasional Paper No. 12, Centre for Urban and Regional Research, University of Manchester.
11. *Regional Shopping Centres in North-West England*—Part II—*A Retail Shopping Model* (1966) Department of Town Planning, University of Manchester.
12. REILLY WJ (1931) *The Law of Retail Gravitation*. Reilly, New York.
13. REYNOLDS RB (1953) A test of the law of retail gravitation. *Jl Marketing* **17**, 273.–277
14. SCOTT P (1970) *Geography and Retailing*. Hutchinson University Library, London.
15. *The* 1971 10% *Sample Census*. The Office of Population Censuses and Surveys, Titchfield, Hants.
16. *The Stores and Shops Retail Directory* (1976) Newman Books, London.

ADDRESS FOR CORRESPONDENCE: *CDT Watson-Gandy Esq, Dept of Management Science, Imperial College, Exhibition Road, London SW7 2BX, UK.*

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 6, 1977. Pergamon Press. Printed in Great Britain

Three-Level Models for Hierarchical Coordination

W DAVIS
University of Illinois

J TALAVAGE
Purdue University

(*Received March 1977; in revised form May 1977*)

Some properties of hierarchical organizations, such as the types of information that may be effectively used for coordination purposes, may be studied by interpretation of mathematical programming decomposition models. Two new models are developed to represent a three-level organization. The models have been implemented on a CDC 6500 computer, and a description is given of their application to a complex transshipment problem. The results of this effort indicate that information regarding 'deviation from goals' is of significant value for coordination of the solution procedure for this type of problem.

INTRODUCTION

A RENEWED interest in the modeling of hierarchical decision-making via the decomposition theories of mathematical programming is emerging. The general concept has its origins in the initial development of decomposition algorithms by Dantzig [1]. He discussed in detail how the Dantzig–Wolfe decomposition algorithm might be applied to model the decision-making process in a two-level organization. Of notable interest was his interpretation of the 'shadow price' (simplex multiplier) as datum for coordinating the individual decision-making subsystems of the hierarchy.

In subsequent years, several new decomposition algorithms have been developed which solve large mathematical programming problems possessing a particular structure. Lasdon [11], Geoffrion [6, 7] and Grinold [8] have recounted many of these decomposition techniques. Often, with the emergence of a new decomposition technique, an attempt was made to describe an organization which mimicked the decomposition strategy. Excellent reviews of the economic

behavioral aspects of the above organizational models have been written by Ruefli [12] and Jennergren [9].

In the mid-sixties however, a new approach to linear programming emerged called goal programming. Using the framework of goal programming and the theory of generalized linear programming, Ruefli [13] developed a new model for a three-level resource allocating organization called the Generalized Goal Decomposition (GGD) algorithm. Unlike previous decomposition models for organizations, the GGD model was evolved strictly to define the structure of a decentralized organization. No overall problem was specified for this organization nor was the decomposition technique upon which it was based given. Later a specification of the GGD model's overall problem was made by Davis and Talavage [2]. They further showed that the GGD algorithm utilized modified restricted masters of generalized linear programs for decisions at the intermediate level, while the decisions at the superordinate and subordinate levels acted as column generators to these restricted master programs.

Freeland and Baker [5] later replaced the superordinate column generator of the GGD model with a restricted master program from the Bender's Partitioning algorithm. They further dropped the lowest level of the hierarchy leaving a two-level hierarchy. Thus the restricted masters of the GGD model were no longer necessary and the second level decisions became a collection of goal programs. Both models use the simplex multiplier as important information for the coordination of the individual decision-makers. The deviations from the prescribed goals generated by the supremal decision-maker for the second level goal programming decision were, however, neglected. Further, both models relied upon cost functions formulated in terms of these deviations alone without regard to cost of the proposed decisions themselves.

Vol. 5
1977

There are possible advantages to be gained in utilizing these deviations as a coordinative input to the goal generating, supremal decision-maker. The incorporation of these deviations leads to the formulation of two new models for the resource allocating organization which are the subject of this paper. The Centralized Goal Decomposition (CGD) model relies only upon the deviations as a coordinative input to the supremal decision-maker while the Hybrid Goal Decomposition (HGD) model uses both the deviations and the associated simplex multipliers to the prescribed goals as coordinative inputs. Both models are formulated for a three-level organization. Further, an overall problem is defined for each algorithm which is identical to that of the GGD model.

The major utility underlying the development of the HGD and CGD models is that both algorithms solve the same overall linear programming problem as the GGD model. Because the GGD model relies solely upon the simplex multiplier as a coordinative input to the supremal goal generator, the set of all three algorithms provides a vehicle for testing the efficacy of different types of coordinating information.

DECISION-MAKING ECHELON

A three-level decision-making echelon, illustrated in Fig. 1, may be described in terms of a central (or supremal) unit with responsibility for coordinating the decisions of M managers. In turn, each of the managing units is responsible for

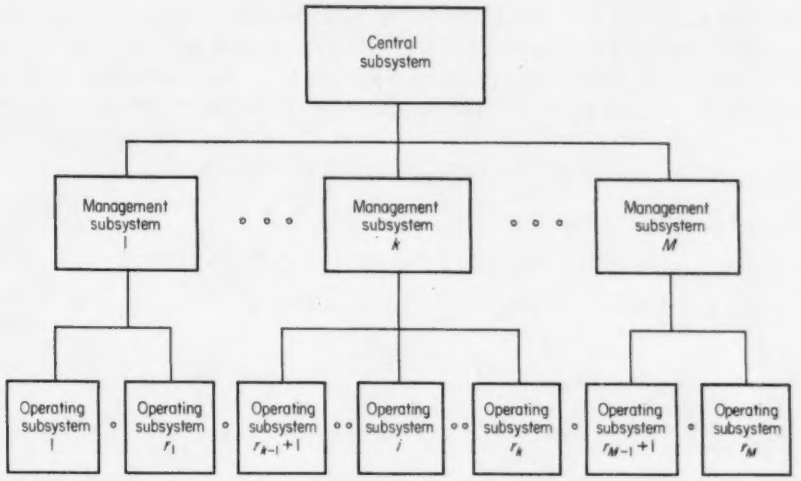


FIG. 1. *Revised echelon for the new algorithm.*

the coordination of a number of operating units. In our notation, each operating subsystem is designated by a single index i. Assuming that there are M managers, we define a series of integers $r_0, r_1, \ldots, r_M$ such that operating sub-systems $(r_{k-1} + 1)$ through r_k report to manager k. For consistency, assuming there are a total of N operating subsystems, we must have $r_0 = 1$ and $r_M = N$.

THE CENTRALIZED GOAL DECOMPOSITION MODEL

The decisions for this model are displayed below by equations (1)–(9). There are two sets of goals to be satisfied by manager k. The first set of goals is the m_k vector $\mathbf{G}_k(t)$, which represents a set of external goals. These goals are generated by the central unit on the t-th iteration to coordinate the action of the managers. The second set of goals, $\mathbf{g}_k$, is a m_k'-vector of internal goals. These goals are constant throughout the calculations and can be regarded as a mechanism used by manager k to coordinate the efforts of the operating sub-systems subject to him.

Corresponding to the external goal vector $\mathbf{G}_k(t)$ are two m_k-vectors, $\mathbf{Y}_k^+(t)$ and $\mathbf{Y}_k^-(t)$, representing positive and negative deviations from $\mathbf{G}_k(t)$ respectively. Similarly, there are two m_k'-vectors, $\mathbf{y}_k^+(t)$ and $\mathbf{y}_k^-(t)$, corresponding to positive and negative deviations from the internal goal vector $\mathbf{g}_k$. Each manager k has, at his disposal, a set of n_i-vector proposals, $\mathbf{X}_i(1), \ldots, \mathbf{X}_i(t)$, generated during the previous iterations by the operating subsystem i, $i = r_{k-1} + 1, \ldots, r_k$. In creating a composite proposal for each operating subsystem i, the manager k can select any convex combination of the aforementioned set, $\mathbf{X}_i(1), \ldots, \mathbf{X}_i(t)$.

Associated with each operating subsystem's proposal is a n_i-vector $\mathbf{c}_i$ such that the cost of a particular proposal $\mathbf{X}_i(j)$ is $\mathbf{c}_i\mathbf{X}_i(j)$. Also associated with the vectors $\mathbf{Y}_k^+(t)$ and $\mathbf{Y}_k^-(t)$ are the m_k-vectors $\mathbf{W}_k^+$ and $\mathbf{W}_k^-$, giving the penalty costs of deviation. Similarly the m_k'-vectors $\mathbf{w}_k^+$ and $\mathbf{w}_k^-$ are associated with $\mathbf{y}_k^+(t)$ and $\mathbf{y}_k^-(t)$. From these costs, equation (4.k) follows.

Central subsystem's problem

$$\text{Min } \Sigma_{k=1}^{M} [\mathbf{C}_k\mathbf{G}_k(t+1) + \mathbf{W}_k^+\mathbf{S}_k^+(t+1) + \mathbf{W}_k^-\mathbf{S}_k^-(t+1)], \tag{1}$$

s.t.

$$\mathbf{G}_k(t+1) + \mathbf{I}_{m_k}\mathbf{S}_k^+(t+1) - \mathbf{I}_{m_k}\mathbf{S}_k^-(t+1) = \mathbf{G}_k(t) + \mathbf{Y}_k^+(t) - \mathbf{Y}_k^-(t) \text{ for } k = 1, \ldots M, \tag{2}$$

$$\Sigma_{k=1}^{M} P_k G_k(t+1) \leqq G_0. \tag{3}$$

Managing subsystem k*'s problem* (k = 1, . . ., M)

$$\text{Min } \Sigma_{i=r_{k-1}+1}^{r_k} \Sigma_{j=1}^{t} \mathbf{c}_i\mathbf{X}_i(j)\lambda_i(j) + \mathbf{W}_k^+\mathbf{Y}_k^+(t) + \mathbf{W}_k^-\mathbf{Y}_k^-(t) + \mathbf{w}_k^+\mathbf{y}_k^+(t) + \mathbf{w}_k^-\mathbf{y}_k^-(t), \tag{4.k}$$

s.t.

$$\Sigma_{i=r_{k-1}+1}^{r_k} \Sigma_{j=1}^{t} \mathbf{B}_i\mathbf{X}_i(j)\lambda_i(j) - \mathbf{I}_{m_k}\mathbf{Y}_k^+(t) + \mathbf{I}_{m_k}\mathbf{Y}_k^-(t) = \mathbf{G}_k(t), \tag{5.k}$$

$$\Sigma_{i=r_{k-1}+1}^{r_k} \Sigma_{j=1}^{t} \mathbf{B}_i'\mathbf{X}_i(j)\lambda_i(j) - \mathbf{I}_{m_k'}y_k^+(t) + I_{m_k'}\mathbf{y}_k^-(t) = \mathbf{g}_k, \tag{6.k}$$

$$\Sigma_{j=1}^{t} \lambda_i(j) = 1, \lambda_i(j) \geqq 0 \text{ for } i = r_{k-1} + 1, \ldots, r_k \tag{7.k}$$

Operating subsystem i*'s problem* (i = 1, . . . , N)

$$\text{Min } (\mathbf{c}_i + \pi_k\mathbf{B}_i + \pi_k'\mathbf{B}_i')\mathbf{X}_i(t+1) \tag{8.i}$$

s.t.

$$\mathbf{D}_i\mathbf{X}_i(t+1) \leqq \mathbf{F}_i \tag{9.i}$$

The managing subsystem's problem is one of generating a composite proposal for each operating subsystem i, $i = r_{k-1} + 1, \ldots, r_k$, satisfying equation (5.k) to (7.k) such that the objective function (4.k) is minimized.

As with the GGD model, the CGD model is also an iterative model. Once the manager k has solved his problem, i.e. to determine the optimum set of $\lambda_i(j)$'s, $\mathbf{Y}_k^+(t)$, $\mathbf{Y}_k^-(t)$, $\mathbf{y}_k^+(t)$ and $\mathbf{y}_k^-(t)$, he has two alternatives to reduce his cost. First he can ask the central subsystem for a new $\mathbf{G}_k(t+1)$, and second he can ask the operating subsystems to generate new proposals. Both alternatives are utilized on each iteration.

Consider the request to the operating subsystems. Associated with the optimal solution to the managing subsystem k's problem is the simplex multiplier

Vol. 5
1977

$\Pi_k(t) \equiv (\pi_k(t), \pi_k'(t), \pi_{0k}(t))$ where $\pi_k(t)$ is a m_k-vector of simplex multipliers associated with equation (5.k), $\pi_k'(t)$ is a m_k'-vector of simplex multipliers associated with equation (6.k) and π_{0k} is the vector $(\pi_{0,r_{k-1}+1}, \ldots, \pi_{0,r_k})$ associated with the constraint (7.k). Suppose a new proposal $\mathbf{X}_i(t+1)$ from the operating subsystem i is to be considered on the next iteration. The relative cost of such a proposal for the current solution is given by $\bar{c}[\mathbf{X}_i(t+1)] = [\mathbf{c}_i + \pi_k(t)\,\mathbf{B}_i + \pi_k'(t)\,\mathbf{B}_i']\,\mathbf{X}_i(t+1) + \pi_{0i}(t)$. Now the manager k desires a new set of proposals $\mathbf{X}_i(t+1)$, $i = r_{k-1}+1, \ldots, r_k$, which will reduce his cost given $\mathbf{G}_k(t)$ and g_k. Thus, he asks each operating subsystem i, $i = r_{k-1}+1, \ldots, r_k$, to generate a new proposal $\mathbf{X}_i(t+1)$ which will reduce the relative cost given above.

Charged with this task, each operating subsystem i begins to minimize its relative cost. Since $\pi_{0i}(t)$ is a constant in the minimization, it is dropped from consideration, giving the objective function (8.i). The operating subsystem i, however, must satisfy the constraint

$$\mathbf{D}_i\mathbf{X}_i(t+1) \leqq \mathbf{F}_i \tag{9.i}$$

in its minimization, where $\mathbf{D}_i$ is an $(l_i \times n_i)$-matrix of linear relations and $\mathbf{F}_i$ is an l_i-vector of stipulations. Assume that equation (9.i) defines a convex polyhedron, say Γ_i. Thus, in the solution of the assigned decision, the vertex of Γ_i which minimizes equation (8.i) will be selected. This is $\mathbf{X}_i(t+1)$ which is returned to manager k for consideration.

Meanwhile manager k has sent his optimal deviations $\mathbf{Y}_k^+(t)$ and $\mathbf{Y}_k^-(t)$ from the current goal $\mathbf{G}_k(t)$ to the central subsystem. Since the central subsystem knows the goal $\mathbf{G}_k(t)$ sent to manager k on the last iteration, the goal that each manager k could meet exactly at his current optimum solution could be determined. This goal is given as $\mathbf{G}_k^*(t) = \mathbf{G}_k(t) + \mathbf{Y}_k^+(t) - \mathbf{Y}_k^-(t)$. The central subsystem is concerned with generating a new set of goals $\mathbf{G}_k(t+1)$, $k = 1, \ldots, M$ which will minimize the deviations from the managing subsystems' current operating levels or goals $\mathbf{G}_k^*(t)$. These deviations are given as $\mathbf{S}_k^+(t+1)$ and $\mathbf{S}_k^-(t+1)$ for $k = 1, \ldots, M$. In selecting the new set of goals, the constraint $\Sigma_{k=1}^M \mathbf{P}_k\mathbf{G}_k(t+1) \leqq \mathbf{G}_0$, where $\mathbf{P}_k$ is a $(m_k \times m_0)$-matrix relating $\mathbf{G}_k(t+1)$ to $\mathbf{G}_0$, a m_0-vector of stipulations or overall goals, must be satisfied. Moreover, in this selection, both the cost of the goals themselves, i.e. $\Sigma_{k=1}^M \mathbf{C}_k\mathbf{G}_k(t+1)$, where $\mathbf{C}_k$ is a m_k-vector of costs, and also the penalty cost of the deviations from the manager k's current operating point $\mathbf{G}_k^*$, i.e. $\Sigma_{k=1}^M [\mathbf{W}_k^+\mathbf{S}_k^+(t+1) + \mathbf{W}_k^-\mathbf{S}_k^-(t+1)]$, must be considered. Combining these costs gives the cost function for the central subsystem as equation (1). Equation (2) relates $\mathbf{S}_k^+(t+1)$, $\mathbf{S}_k^-(t+1)$ and $\mathbf{G}_k(t+1)$ to $\mathbf{G}_k^*$. In solving this problem we can be assured that

$$\Sigma_{k=1}^M [\mathbf{C}_k\mathbf{G}_k(t+1) + \mathbf{W}_k^+\mathbf{S}_k^+(t+1) + \mathbf{W}_k^-\mathbf{S}_k^-(t+1)] \leqq \Sigma_{k=1}^M [\mathbf{C}_k\mathbf{G}_k(t) + \mathbf{W}_k^+\mathbf{Y}_k^+(t) + \mathbf{W}_k^-\mathbf{Y}_k^-(t)]$$

since the solution $\mathbf{S}_k^+(t+1) = \mathbf{Y}_k^+(t)$, $\mathbf{S}_k^-(t+1) = \mathbf{Y}_k^-(t)$ and $\mathbf{G}_k(t+1) = \mathbf{G}_k(t)$, $k = 1, \ldots, M$, satisfies equations (2) and (3).

Having generated the optimal set of goals $\mathbf{G}_k(t+1)$, $k = 1, \ldots, M$, the central subsystem sends each goal $\mathbf{G}_k(t+1)$ to manager k. Time is then incremented from t to $t+1$, and the process is repeated.

THE HYBRID GOAL DECOMPOSITION MODEL

The Hybrid Goal Decomposition (HGD) model utilizes concepts from both Ruefli's GGD model and the CGD model. The equations for the managing and operating subsystems are identical to those of the CGD model.

The difference in the algorithms lies in the interaction between the central unit and the management subsystems, as shown by equations (10)–(15) below. To assist the central unit in its decision, each manager supplies the feedback information of $\mathbf{Y}_k^+(t)$, $\mathbf{Y}_k^-(t)$, *and* the simplex multiplier, $\pi_k(t)$.

Central subsystem's problem

Decision I

$$\text{Min} \quad \Sigma_{k=1}^{M} (\mathbf{C}_k - \pi_k(t))\mathbf{G}_k'(t+1) \tag{10}$$

s.t.

$$\Sigma_{k=1}^{M} \mathbf{P}_k\mathbf{G}_k'(t+1) \leqq \mathbf{G}_0 \tag{11}$$

Decision II

$$\text{Min} \quad \Sigma_{j=1}^{t+1} \Sigma_{k=1}^{M} \mathbf{C}_k\mathbf{G}_k'(t)\beta(j) + \Sigma_{k=1}^{M} [\mathbf{W}_k^+\mathbf{S}_k^+(t+1) + \mathbf{W}_k^-\mathbf{S}_k^-(t+1)] \tag{12}$$

s.t.

$$\Sigma_{j=1}^{t+1} \mathbf{G}_k'(j)\beta(j) + \mathbf{I}_{m_k}\mathbf{S}_k^+(t+1) - \mathbf{I}_{m_k}\mathbf{S}_k^-(t+1) = \mathbf{G}_k(t) + \mathbf{Y}_k^+(t) - \mathbf{Y}_k^-(t)$$

$$\text{for } k = 1, \ldots, M \tag{13}$$

$$\Sigma_{j=1}^{t+1} \beta(j) = 1, \beta(j) \geqq 0 \tag{14}$$

where

$$\mathbf{G}_k(t+1) = \Sigma_{j=1}^{t+1} \mathbf{G}_k'(j)\beta(j) \text{ for } k = 1, \ldots, M \tag{15}$$

In the HGD model, the central subsystem has two problems to solve. The first decision is concerned with finding the vertex of the convex polytope given by equation (11) as $\Sigma_{k=1}^{M} \mathbf{P}_k\mathbf{G}_k'(t+1) \leqq \mathbf{G}_0$ which best minimizes the marginal cost of the vertex with respect to the manager's current solutions. Since the central subsystem has $\pi_k(t)$ for each of the managing subsystems, it can determine this marginal cost for any $\mathbf{G}_k'(t+1)$ as

$$\bar{\mathbf{C}}(\mathbf{G}_k'(t+1)) = (\mathbf{C}_k - \pi_k(t))\mathbf{G}_k'(t+1).$$

However, since $\mathbf{G}_k'(t+1)$ must be generated as an element of the matrix $[\mathbf{G}_1'(t+1), \ldots, \mathbf{G}_M'(t+1)]$, it is decided to minimize the sum of the marginal cost for each $\mathbf{G}_k'(t+1)$, given by equation (10) over the convex polytope given by equation (11). This decision is identical to that of the central subsystem in Ruefli's GGD model. Rather than returning this vertex to the managing subsystems as their new goal $\mathbf{G}_k(t+1)$, which is the case in the GGD model, the central subsystem then uses past knowledge gained in its decision-making.

Vol. 5
1977

It begins first by computing the goal that the managers could meet exactly at their present solution, namely $\mathbf{G}_k^*(t) = \mathbf{G}_k(t) + \mathbf{Y}_k^+(t) - \mathbf{Y}_k^-(t)$. Taking the vertices of equation (11), generated during past iterations, along with the new vertex, $[\mathbf{G}_1'(t+1), \ldots, \mathbf{G}_M'(t+1)]$, convex combination of the vertices for $\mathbf{G}_k(t+1)$ are examined, i.e. $\mathbf{G}_k(t+1) = \Sigma_{j=1}^{t+1} \mathbf{G}_k'(j)\beta(j)$ such that $\Sigma_{j=1}^{t+1} \beta(j) = 1$, $\beta(j) \geqq 0$. The particular convex combination is chosen such that the actual cost of the goals, $\Sigma_{k=1}^{M} \mathbf{C}_k\mathbf{G}_k(t+1)$ plus the cost of deviation from $\mathbf{G}_k^*(t)$ is minimized. This problem is posed by equations (12)–(14). The decision itself resembles that of the central subsystem in the CGD model. Once the optimal $\mathbf{G}_k(t+1)$, given by equation (15), has been selected, it is returned to the managing subsystem k and the iterative algorithm continues as described in the CGD model.

SOLUTION PROPERTIES

In [3], an overall problem is defined for both the CGD and HGD model as well as for the GGD model. Further it can be shown that the CGD and HGD models generate feasible solutions to this problem on every iteration. If one makes the following assumption, convergence can be shown.

Assumption: if, on any iteration, the decision for a given decision-making subsystem on the previous iteration is an alternate optimum solution to his present problem, the solution from the previous iteration is repeated. This assumption results in a sequence for the performance criterion of the organization on iteration t, $z(t)$, such that

$$z(1) > \cdots > z(t) = z(t+1) = \cdots = z(t+T) \text{ for all } T > 0.$$

Thus, the model has converged.

If the above assumption is removed, it can be shown that $z(t+1) \leq z(t)$ for all t. However, $z(t+1) = z(t)$ does not necessarily imply that $z(t+T) = z(t)$ for all $t > 0$. In the following application of the CGD and HGD models, the above assumption has been employed.

COMPUTATIONAL EXPERIENCE

Both the CGD and HGD algorithms have been applied to several examples. Perhaps the most extensive model used to date is a model of a three-level transshipment organization shown in Fig. 2. As can be seen, the lowest level of the organization is the transshipment network where each node acts as a decision maker. Authority over subsets of these nodes has been delegated to the managers on the second level.

These nodes, acting as decision-makers, know their needs and the capacities of the branches converging upon them. On this basis, they choose what they

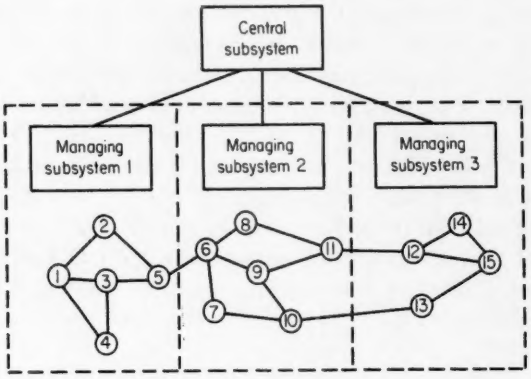


FIG. 2. *Organizational structure for transshipment example.*

perceive to be an efficient operating proposal, irrespective of the other nodes. A substantial coordination problem arises. The problem is to insure that what one nodes expects to receive from other nodes is the same as what the other nodes expect to ship to the first. If all these nodes are under the same manager k, i.e. elements of the subset A_k, the coordination can be handled internally on the second level.

However, the nodes can and do interact with nodes under different managers. In this situation, a mechanism for coordinating the actions of the managers is needed. This gives rise to the need for the central manager who has authority over the other managers and represents the highest level in our organization. The central manager has also been given the authority for establishing the level of production at the source nodes in the transshipment network so as to guarantee an overall material balance for the total network and for each manager.

The actual equations necessary to effect this coordination are too extensive to be included in this paper. (The interested reader is referred to [3].) The remainder of this paper will focus upon the results attained when the CGD and HGD algorithms were applied to the particular transshipment organization depicted in Figs. 2 and 3.

One should note that nodes 1, 9, and 15 are sources. Two of the source nodes were deliberately placed at the extremes of the network to create a need for substantial transshipment through intervening nodes. Furthermore, to assure the necessity for transshipment, the nodal demands and the capacity for production at the source nodes were set so as to require Manager 2 to import quantities from nodes under both Manager 1 and 3; the production limits for source nodes 1, 9, and 15 were set at 575, 500 and 400, respectively.

The actual optimum solution for the transshipment network is given in Fig. 4 and the cost is \$36,280. Both the CGD and HGD models converged to final

Vol. 5
1977

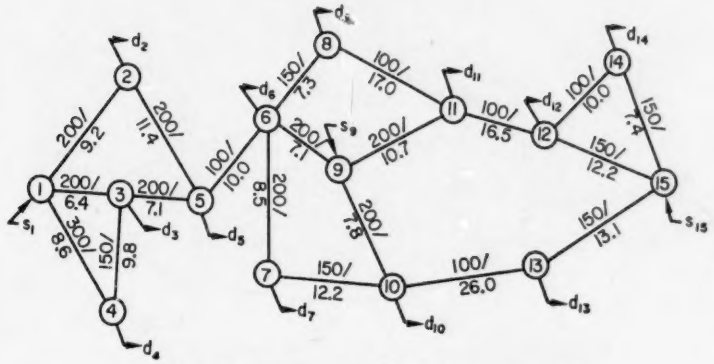


FIG. 3. *Detailed network for transshipment example.*

solutions in approx. 40 iterations. Table 1 lists the network flows generated by each model in relation to the optimum solution. The final costs of these solutions are \$37,585 and \$37,109 for the CGD and HGD models respectively. This

TABLE 1.

	OPT	CGD	HGD
X_{12}	175	198·7	200
X_{13}	200	175·3	175
X_{14}	200	200·4	200
X_{25}	125	148·7	150
X_{35}	125	101·3	100
X_{43}	0·0	0·9	0
X_{56}	100	100	100
X_{67}	200	165	200
X_{68}	10	8·4*	10
X_{96}	200	163·4	200
$X_{9,10}$	165	200	165
$X_{9,11}$	135	136·6*	135
$X_{10,7}$	50	85	50
$X_{11,8}$	60	61·7*	60
$X_{12,11}$	100	100	100
$X_{13,10}$	35	35	35
$X_{14,12}$	40	54·7	40
$X_{15,12}$	150	135·3*	150
$X_{15,13}$	75	75	75*
$X_{15,14}$	120	134·7*	120
Z	36280	37585·5	37109·1

* The flow given is the computed net flow along the arc.

confirms that the CGD and HGD models may not converge to the optimum solution. It is important to note however, that both the CGD and HGD models did generate feasible solutions to the organization's transshipment network.

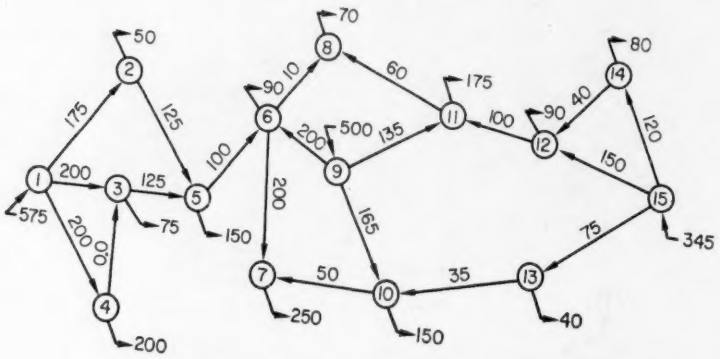


FIG. 4. *Network display of optimum solution for transshipment example.*

CONCLUSIONS

Before drawing any conclusions, perhaps one should look at the development of the CGD and HGD models. Neither model was developed solely as a new decomposition algorithm for solving a large linear programming problem. Rather both were designed to model the behavior of a particular goal-seeking organization. Of special interest to the authors was the ability of each model to effect the solution of the organization's problem *given* the organizational structure.

Ruefli [12], in his section on behavioral models, notes that many new models do not tend to the optimum solution of the organization's problem. Instead the classical optimization approach has been displaced by satisficing or a multiple goal approach. In his GGD model [13], this approach is reflected in that each managing subsystem has its own objective function which is never considered in relation to those of the other managing subsystems.

The CGD and HGD models are extensions of the GGD model. However, we have attempted to retain the structure of the organization proposed by Ruefli while replacing the satisficing concept by an attempt to optimize the organization's overall problem. We have shown that both the CGD and HGD models do converge to a limitpoint, but not necessarily to the optimum solution. (Note that Ruefli's GGD when applied to this problem did not converge [3].) This resulted from the fact that in the CGD and HGD model, no subsystem solves the master program resulting from a classical decomposition of the organization's problem. Instead the master program is decoupled into a set of subproblems for the central and managing subsystems using a directional search approach.

The usefulness of this type of research has been discussed by Kornai [10], Ruefli [12], and others. Kornai notes that the computational efficiency of such

Vol. 5
1977

models is not the real issue. Rather that the development of such models will hopefully lead to a better understanding of organizations and provide a mathematical theory to aid in their design. To this end, Kornai asks how information flows between the subysystems might be designed so that the number of interactions (iterations) between them can be minimized? This development of the CGD and HGD models has addressed that issue. In particular, the HGD model incorporates the concept of passing the simplex multipliers *and* the deviation vectors from the managing subsystems as input to the central subsystem's decision (as opposed to only the simplex multiplier vectors for the GGD model and only the deviation vectors for the CGD model). The results indicate the importance of the deviation vector as an input to the central subsystem's decision, whereas the simplex multiplier as an additional input to this decision in the HGD model is of questionable value. Its performance was only slightly better than the CGD model for this example and others [3]. Indeed there is some evidence here to demonstrate that the simplex multiplier impairs the effective coordination of the operating subsystems' decision.

In Table 1, several flows for both the HGD and CGD model have asterisks attached. These flows are net flows; i.e. in the final solutions, flows exists in both directions. It would seem that the back flows along an arc could easily be removed. The values assigned to the simplex multiplier however did not allow this removal due to the penalty cost functions, $\mathbf{W}_k{}^+\mathbf{Y}_k{}^+(t)$ and $\mathbf{W}_k{}^-\mathbf{Y}_k{}^-(t)$. Looking at the managing subsystem k's problem, he has a set of vector proposals $\mathbf{X}_i(j), j = 1, \ldots, t$, from each node $i, i = r_{k-1} + 1, \ldots, r_k$. The actual number of proposals among this set is usually less than the number of constraints for the managing subsystem k's decision. Furthermore the generation of these proposals by the operating subsystems and subsequent multiplication of $\mathbf{X}_i(t)$ by matrices $\mathbf{B}_i$ and $\mathbf{B}_i'$ leads to round-off errors. The net result is that components of the deviation vector must remain in the solution at a very small value. From a conceptual standpoint all the equality constraints have been met. However, the simplex multiplier reflects only that the constraint is not being met, i.e. the deviation is positive. Since the cost attached to the deviation vectors in the penalty cost functions is the largest and this cost influences the simplex multiplier, we must conclude that the coordination between the operating subsystems and their managing subsystem has been impaired.

Current research is directed towards eliminating the simplex multiplier from the decisions of all subsystems. This has resulted in a generalized model which allows the inclusion of linear, nonlinear and/or integer constraints into the decisions. Computational testing of the proposed model is now being studied.

In summary, two new models for organizational coordination have been developed, implemented, and applied to a complex transshipment problem. The results of this effort indicate that information regarding 'deviation from goals' is of significant value for coordination, and appears to eclipse the corresponding value of information regarding 'prices' for this type of problem.

REFERENCES

1. DANTZIG GB (1963) *Linear Programming and Extensions*, pp. 448–469. Princeton University.
2. DAVIS WJ and TALAVAGE JJ (1974) Goal programming: Dantzig–Wolfe in disguise? Presented as a session paper at ORSA/TIMS Puerto Rico Meeting.
3. DAVIS WJ (1975) *Three-Level Echelon Models for Organizational Coordination*. Ph.D. dissertation, Purdue University.
4. DAVIS WJ and TALAVAGE JJ (1976) Programming for Organizational Coordination. *Proceedings of Southeastern AIDS Conference*.
5. FREELAND JR and BAKER RN (1975) Goal partitioning in a hierarchical organization. *Omega* 3(6), 673–688.
6. GEOFFRION AM (1970) Elements of large-scale mathematical programming. *Mgmt Sci.* **16**(11), 652–691.
7. GEOFFRION AM (1971) Large-scale linear and nonlinear programming, in *Optimization Methods for Large-Scale Systems* (Ed. WESMER DA) pp. 47–74. McGraw-Hill, New York.
8. GRINOLD RC (1972) Steepest ascent for large-scale linear programs. *SIAM Rev.* **14**(3), 447- 464.
9. JENNERGREN LP (1971) *Studies in the Mathematical Theory of Decentralized Resource-Allocation*. Ph.D. Dissertation, Stanford University.
10. KORNAI J (1970) Thoughts on Multi-Level Planning Systems, in *Multi-Level Planning—Case Studies in Mexico* (Eds. GORLUK E. and MANNE A.) pp. 522–551. North Holland–American Elsevier, New York.
11. LADSON LS (1970) *Optimization Theory for Large Systems*. Macmillan, New York.
12. RUEFLI TW (1969) *Planning in Decentralized Organizations*. Ph.D. Dissertation, Carnegie-Mellon University.
13. RUEFLI TW (1971) A generalized goal decomposition model. *Mgmt Sci.* **17**(8), 505–518.

ADDRESS FOR CORRESPONDENCE: *Joseph Talavage Esq, School of Industrial Engineering, Purdue University, Grissom Hall, West Lafayette, IN* 47907, *USA.*

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 6, 1977. Pergamon Press. Printed in Great Britain

Dynamic Booking Rules for Flights with an Intermediate Stop

SP LADANY
Ben-Gurion University

DN BEDI
The City University of New York

(*Received May 1976; in revised form July 1977*)

A feasible model has been developed to find the optimal decision rule for allocating seats to passengers seeking to fly on the long or short span of a flight with an intermediate stop from which no boarding rights exist to take passengers to the terminal stop. The model is feasible from the point of view of estimating the data and of computer calculation and memory capacity and is implementable under real life conditions. A computer program was written and a short sample problem was calculated to show the differences in computation time and memory requirements vs the loss of revenue in comparison with a complicated and thus unfeasible higher order model.

INTRODUCTION

MANY airlines schedule their flights in a manner that the airplane has an intermediate stop between the starting and the terminal points. The intermediate stop is short and it is used for refuelling and for discharge of passengers flying just to the intermediate stop, or for boarding of new passengers. However, in international flights, the rights of landing for refuelling, of discharging and/or boarding of passengers are regarded by the various countries as national economic assets, and therefore they are chartered and strictly enforced. The common situation for national airlines is to obtain—reciprocally—rights to service the flow of passengers between their home bases, but they are limited to landing and discharging of passengers at other foreign cities. Thus, when a flight continues to another destination, in the same foreign country or somewhere else abroad, the airline is prevented from boarding passengers at the intermediate stop in order to fly them to the flight's terminal destination. Such restrictions are enforced not only in international flights, where it is almost dominantly practiced in the cross-Atlantic market, but national aeronautical boards in various countries

regulate their air transportation by granting such constrained charters to the various domestic and international airlines over the various routes.

Hence, once the route is known, the airline is faced with the problem of deciding in advance on the allocation of seats. It may allocate the seats either to passengers wishing to fly from the starting point to the terminal point or to those passengers wishing to fly just to the intermediate stop. Clearly, a passenger flown the whole route contributes more to the profit per flight than the passenger who pays to be flown only to the intermediate stop. Thus, under deterministic conditions, there would be no problem on booking passengers on the long or the short segments of the flight: all the available passengers seeking to be flown on the long segments will be assigned seats, if available, and the remaining seats will be at the disposal of the passengers seeking to be flown to the intermediate stop. However, the problem is complicated because of the probabilistic—stochastic—nature of the demand and cancellations which lead to possible overbooking policies, and to the fact that last moment unbooked passenger demand (waitlist and late arrival) might be created and such demand is also distributed probabilistically. To exemplify the problem, consider a case of a scheduled flight from Athens to New York with an intermediate stop in London. During the booking period in the hours, days and weeks prior to the flight when there are still unbooked seats available, it must be decided whether to book a passenger requesting to be flown from Athens only as far as London. If booked, and if the passenger will not cancel his reservation, he will generate a revenue, although this revenue is less than what would have been obtainable from assigning the seat to an Athens–NY passenger. Alternatively, the airline may refuse the booking request (and waitlist), claiming an already fully booked flight, gambling on the hope of finding a passenger willing to fly from Athens to New York (and assuming that he will not cancel his reservation) or of finding later on at least another passenger willing to fly just to London, and risking the loss of the revenue from the passenger dealt with.

Vol. 5
1977

The problem of airline cancellations and their influence on bookings was investigated by Martinez and Sanchez [4]. Their extensive empirical study is the basis for the assumptions concerning the distribution of cancellations as used in this paper. Griffiths and Taylor [1] had developed a complicated mathematical model for BOAC for booking passengers to different flight sectors; they did not consider overbooking and late arrivals, and their model has not been implemented by BOAC or by any other airline. The problem of reservations with overbookings and standbys was treated by Rothstein [5], but in that work no consideration was given to the problem of a flight with intermediate stops and the resultant allocation of seats to passengers for the various segments. Hersh and Ladany [2] developed a model which attempts to find an optimal decision rule for allocating seats to passengers for the full and the short flight spans considering the possibility of overbooking cancellations and standbys. Moreover, their model involves a continuous dynamic Bayesian reassessment of future

demand probabilities. Ladany and Bedi [3] have further shown that the model developed in [2] can be adjusted to analyze the advantage of separated vs continuous boarding procedures.

The Hersh and Ladany [2] model, though theoretically workable and though considering all the behaviourial facets of the reservation-booking problem has serious weaknesses which hamper, decrease, and in many cases prevent completely its implementation. First, it requires the estimation of a huge amount of demand distributions: the distribution of demand as a function of time prior to flight for various final demand levels, separately for the long and for the short spans. From the discussion of Rothstein and Etschmaier [6, p.170] on the difficulties of estimating the passenger demand distribution for use in mathematical booking models, it is evident that for [2] the real life difficulties of data estimation become enormous and intolerable, Second. the structural incorporation in the model of the Bayesian reassessment of probabilities requires the use of four variables, and when each demand variable might take on 100 different values, the problem rises to having the impossible and intolerable dimension of $(100 \times 50.5)^2 = 25{,}502{,}500$. Existing computers cannot compute at all problems of such dimension in an economic manner. Neither can they store the results in memory for on-line booking purposes. Further it is evident that the huge dimension of the problem prevents implementation of the method, if at all possible, to a noncomputerized booking system which uses for booking decisions pre-calculated tables.

In the present paper a model has been developed to find an optimal decision rule for allocating seats to passengers for the various segments, and which would be feasible from the point of view of estimating the data and of computer calculations and memory capacity. To achieve this aim, the model of [2] was restructured and the number of demand variables was reduced from four to two. Such a non-Bayesian model, in considering a range of 100 values for each variable would require just a total dimension of 10,000, and would be feasible for implementation. The paper intends to point out the advantage and feasibility of the proposed model, relatively to Hersh and Ladany's model by showing the differences in computation time versus loss of revenue, for a simple case which can be handled by both models.

THE OBJECTIVE

The objective for the airline is to determine an operating policy for allocation of seats to passengers flying full and partial spans such that the expected total contribution to profit for a flight is maximum. Since demands and cancellations for the full span and partial segment are stochastic variables depending on the time prior to the flight departure it becomes necessary to derive the allocation policy as a sequential decision process. This follows the objective stated in [2].

The booking period prior to flight is divided into t decision periods (not necessarily of equal lengths) as for example the last 2 hours ($t = 1$), the previous day ($t = 2$), the previous week ($t = 3$), etc. It seems reasonable to assume that at an early date prior to flight, the interval between decisions should be long, for instance on the order of magnitude of weeks or months, while hourly decisions may be appropriate during the last day prior to flight, but order does not have to be ranked consistently greater.

At the start of each period, t, a decision must be made about the maximum number of additional bookings, $k(t)$ and $h(t)$ that can be allowed during period t for the full and short span, respectively. This decision must be based upon:

(a) The capacity of the aircraft, C.

(b) The contribution to profit per flight for both the full span and short span. This is denoted by a and b, respectively.

(c) The overbooking penalty. This is π_a for passengers holding a reservation for a seat on the full span who must be refused boarding due to a lack of space on the aircraft at the time of departure. The overbooking penalty for short span passengers will be denoted by π_b.

(d) The probabilistic pattern of future demands, cancellations and no-shows for seats on the two spans.

(e) The distributions of late, unbooked passengers (including waitlist passengers) arriving at time of boarding (departure) and desiring a seat on the one or the other span.

(f) The number of passengers already booked on the full and partial span, n and q, respectively.

Vol. 5
1977

THE MODEL

Assume that for a particular flight, we are considering a time t periods before that flight departure, at which n passengers for the full pan and q passengers on the partial span have already been booked.

Then, the maximum total expected contribution to profit at period t for the flight can be denoted by $V_t(n, q)$. Assuming that the allotment of k and h seats is made for the period t before departure then the maximum total expected-contribution to profit should be expressed as a recursion formula: it has to depend on the maximum total expected contribution to profit in the previous period and on the changes that might occur in it during a new decision perod due to cancellations by passengers booked until the end of the last period and due to new requests for booking. Thus:

$D_t(i)$ and $E_t(j)$ are the probabilities of i requests for seats on the full span and j requests for seats on the partial span, respectively, during period t prior to the booking date.

$R_t(n, l)$ and $S_t(q, g)$ are the probabilities of cancellations of l reservations out

of n previous bookings and of g cancellations out of q previous bookings during period t for seats on the long span and seats on the partial span, respectively.

Hence the maximum total expected contribution to profit per flight at period t can be shown to be equal to

$$V_t(n, q) = \underset{k, h}{\text{Max}} \left\{ \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} D_t(i) E_t(j) \sum_{l=0}^{n} \sum_{g=0}^{q} V_{t-1}(n - l + i^*, q - g + j^*) R_t(n, l) S_t(q, g) \right\}, \quad (1)$$

where $i^* = \text{Min}(i, k)$, $j^* = \text{Min}(j, h)$.

The initial conditions required for the solution of the recursive relationship of equation (1) are based on the following policy:

1. The contribution to profit of a seat allocated to a passenger requesting to be flown the full span is greater than that allocated to a passenger requesting to be flown the partial span, and further that the overbooking penalty of the former is greater than that of a partial span passenger. Therefore, it becomes obvious that preference in allocating seats at time of departure should always be given to the booked passenger requesting a seat on the full span.
2. Priority is given to overbooked passengers over unbooked passengers.
3. Priority is given to unbooked passengers requesting a seat on the full span over unbooked passengers requesting a seat on the partial span.

The initial conditions of equation (1) for $t = 0$ depend on the number of booked passengers of both types at the time of departure, and on their numerical relationship to the airplane's seat capacity. Thus, the initial conditions can be separated into three mutually exclusive cases:

(I) If the number of booked passengers for the full span at time if boarding is bigger than the capacity of the airplane, than the expected contribution to profit per flight would be the contribution to profit per passenger booked for the full span times the seat capacity of the airplane, less the number of overbooked passengers in the full span times the overbooking penalty per passenger on the full span, and less the number of passengers booked on the short span times the overbooking penalty per passenger on the short span.

(II) If at time of boarding the number of booked passengers on the full span is not more than the capacity of the airplane, but the combined number of booked passengers on both spans is at least as the seat capacity of the airplane, then the expected contribution to profit per flight consists of the following terms: the contribution to profit from *all* the passengers booked for the full span, plus the contribution to profit per passenger booked for the short span times the airplane capacity not alloted to full span passengers, and less the overbooking penalty per passenger on the short span times the number of overbooked passengers on the short span.

(III) If at time of boarding the combined number of passengers booked on both spans is less than the capacity of the airplane, then the expected contribution to profit per flight will consist of the contribution to profit from all the booked

passengers on both spans, plus terms consisting of the expected contribution to profit from filling up the unbooked seat capacity by the available standby or waitlist passengers available at the time of departure for the full span, and from filling up the remaining unused seat capacity, in that order, by the available standby or waitlist passengers for the short span.

In case booked passengers are admitted to the airplane at time of boarding according to their sequential order of arrival, and not in a separated manner as implied by policy statement No. 1 above and by the resulting equations explained under (I), (II) and (III), the initial conditions of $V_o(n, q)$ become as stated in [3]; i.e. (I) and (II) should be replaced by the following:

If the combined number of booked passengers on both spans is at least as the capacity of the airplane, then the expected contribution to profit per flight would consist of:

(i) the contribution to profit per passenger booked for the full span times the number of passengers booked for the full span plus the contribution to profit per passenger booked for the short span times the number of passengers booked for the short span, both of them multiplied by the proportion of the passengers seated out of the total booked passengers and less

(ii) the number of passengers booked for the full span times the overbooking penalty per overbooked full span passenger, plus the number of passengers booked for the short span times the overbooking penalty per overbooked short span passenger, both of them multiplied by the proportion of unseated passengers out of the total booked passengers.

Vol. 5
1977

THE DATA

The probability distribution of cancelled bookings (including 'no shows') for seats on the full span, $R_t(n, l)$, may be evaluated by assuming that each booked passenger has a constant probability r_t at the tth decision period prior to departure of cancelling his reservation. Therefore, the cancellation of l reservations from a total of n bookings during the tth decision period is binomially distributed, i.e.

$$R_t(n, l) = \binom{n}{l} [r_t]^l [1 - r_t]^{n-l}. \tag{2}$$

Similarly, if each booked passenger on the partial span has a constant probability u_t at the tth decision period prior to departure of cancelling his reservation, the corresponding binomial probability of cancelling G reservations from a total of q bookings during the tth decision period is

$$S_t(q, g) = \binom{q}{g} [u_t]^g [1 - u_t]^{q-g}. \tag{3}$$

The assumed values of r_t and u_t, taken from [2], are given in Table 1.

TABLE 1. PROBABILITY OF A PASSENGER TO CANCEL HIS RESERVATION

Decision period t	Probability of cancellation	
	Full span r_t	Partial span u_t
0	0	0
1	0·03	0·04
2	0·05	0·03
3	0·05	0·05
4	0·02	0·02
5	0	0

The probability distributions of the number of requests for seats on the full span, $D_t(i)$, and the number of requests for seats on the partial span, $E_t(j)$, during the tth decision period prior to departure, can be evaluated from analyzing the history of requests for reservations during the time prior to departure included in the tth decision period. It is possible that the Poisson or the normal distributions might allow a good representation of the empirical data, but a non-fitted empirical distribution could be used as well without complicating the utilization of the model. In order to allow valid comparison of the proposed model with the result of [2], the probability distributions of $D_t(i)$ and $E_t(j)$ have been derived from the data assumed in [2] and are presented in Table 2. (See Appendix 1.)

RESULTS AND CONCLUSIONS

For an airplane capacity of $C = 6$, contributions to profits of $a = 100$, $b = 40$ and overbooking penalty of $\pi_a = 1000$ and $\pi_b = 600$, the expected contribution to profit per flight—obtainable by using the suggested policy—was calculated. It was derived by the use of the recursion formula, equation (1), and from the relationship that

$$E[V_t(n, q)] = \sum_{\text{all } n} \sum_{\text{all } q} V_t(n, q)P(n)P(q) = V_{t\max}(n = o, q = o),$$

where $t_{\max}$ is the last decision period, prior to which no requests for bookings are received.

For the sequential boarding policy, it was found that $E[V_t(n, q)] = 398.05$. This represents just a decrease of 4.83% from the corresponding expected

contribution to profit per flight of 418.26 theoretically obtained from utilizing Hersh and Ladany's model with the same operational data. Such a decrease in contribution to profit seems reasonable and acceptable in view of the facts that (1) the computer computation time with an inefficient primary program was cut from 12 seconds to below 0.1 sec; (2) the computer core requirements are very much reduced, and (3) the output of the program for use of the booking desk will consists of a short list containing—in the present case—just $8 \times 5 = 40$ rows for each decision period t, as opposed to a list with $(8 + 1)8/2 \times (5+1)5/2 = 540$ entries for the older model. These computational advantages become more pronounced with the increase of the seat capacity, and at a certain level the proposed model is not just the economically feasible one but also the only method which allows—computationally and administratively—handling the problem and implementing the results.

The sequential decision process which has been developed provides the airline reservation system with a sequential decision tool for allocating the seats on a flight to passengers desiring to be flown the full or the partial span of the flight. The approach offers a series of decisions where each decision uniquely applies to a specific calendar date, i.e. for February 1, for February 2, etc. The optimal

Vol. 5
1977

TABLE 2. PROBABILITY DISTRIBUTIONS OF DEMAND FOR SEATS ON THE FULL AND PARTIAL SPANS*

	$D_t(i)$				
i	$t = 1$	2	3	4	5
0	0·215158	0·316879	0·476349	0·623298	0·773809
1	0·338482	0·355039	0·342066	0·282881	0·185671
2	0·280238	0·224901	0·137463	0·076742	0·035226
3	0·127022	0·082709	0·036798	0·014874	0·004823
4	0·033528	0·028002	0·006552	0·002021	0·000445
5	0·005082	0·002299	0·000724	0·000175	0·000025
6	0·000467	0·000165	0·000045	0·000008	0·000001
7	0·000022	0·000006	0·000001	0	0

	$E_t(j)$				
j	$t = 1$	2	3	4	5
0	0·347210	0·546240	0·767785	0·849000	0·898085
1	0·447560	0·348387	0·197985	0·133447	0·094235
2	0·184060	0·094894	0·030835	0·016154	0·007285
3	0·020360	0·010087	0·003235	0·001347	0·000385
4	0·000810	0·000391	0·000150	0·000051	0·000110

*For the derivation, see Appendix 1 and [2].

decision for a specific date would be obtained by solving equation (1) from a table which would state—at various time intervals prior to departure, and for all logically possible combinations of values of n and q—the values of k and h. The result would be that the total expected contribution to profit from flight bookings is maximum. An example of the printout of the program for the $t = 2$ decision period is given in Table 3. Similar tables have been obtained for all the other decision periods for the same flight.

TABLE 3. PRINTOUT OF THE PROGRAM FOR $t = 2$

n	q	k	h	n	q	K	h
0	0	4	2	4	0	0	0
0	1	4	1	4	1	0	0
0	2	4	0	4	2	0	0
0	3	3	0	4	3	0	0
0	4	2	0	4	4	0	0
1	0	3	1	5	0	0	0
1	1	3	0	5	1	0	0
1	2	3	0	5	2	0	0
1	3	2	0	5	3	0	0
1	4	1	0	5	4	0	0
2	0	2	1	6	0	0	0
2	1	2	0	6	1	0	0
2	2	2	0	6	2	0	0
2	3	1	0	6	3	1	1
2	4	0	0	6	4	1	0
3	0	1	0	7	0	0	0
3	1	1	0	7	1	0	0
3	2	1	0	7	2	0	0
3	3	0	0	7	3	0	1
3	4	0	0	7	4	0	0

It should be noted that flights in various seasons, or various days in a week, or even at different departure times might requre the estimation of the appropriate distributions of $D_t(i)$, $E_t(j)$, $R_t(n, l)$ and $S_t(q, g)$ and the calculation and application of the corresponding different booking policies.

Application of the process would require the acquisition of a relatively small amount of data from the past reservation history. Grouping of demand data and a reduction in the number of decision periods may be desirable to save computation and to reduce the dimension of the problem in cases where the reservation horizon is extremely long or where the number of requests for reservations at the various decision periods or the airplane's seat capacity is extremely large.

The derived optimal booking policy decision for various flights can be calculated by computer, printed as tables and kept with the airline's chief booking clerk. Alternatively, for airlines that use on-line computerized reservation systems, the table could be kept in the memory of the computer, and a decision

to book or refuse booking a new passenger requesting reservation for a given flight could be revealed to the booking clerk by the computer.

The optimal reservation policy for return flights from a terminal point to the home base with an intermediate stop-over, where boarding rights exist just to fly passengers from the terminal point of the home base or from the intermediate point to the home base, but not from the terminal point to the intermediate point, can be derived similarly to the method presented by adjusting the developed model to the new boarding right conditions.

REFERENCES

1. GRIFFITHS B and TAYLOR J (1967) Mathematical formulation of the sector control problem. *Proceedings VII AGIFORS Meeting*, Vol. 7.
2. HERSH M and LADANY SP Optimal seat allocation for flights with intermediate stops. *Comput. Ops Res.* (in press).
3. LADANY SP and BEDI DN (1975) Analysis of different boarding procedures. *Management Science Applications to Leisure-Time Operations* (Ed. LADANY SP) pp 197–213. North-Holland, Amsterdam.
4. MARTINEZ R and SANCHEZ M (1970) Automatic booking level control. *Proceedings X AGIFORS Meeting*, Vol. 10.
5. ROTHSTEIN M (1971) An airline overbooking model. *Transportation Sci.* 5(2). 180–192.
6. ROTHSTEIN M and ETSCHMAIER MM (1974) Operations research in the management of the airlines. *Omega* 2(2), 157–179.

Vol. 5
1977

ADDRESS FOR CORRESPONDENCE: *Professor SP Ladany, Dept of Industrial Engineering and Management, Ben-Gurion University of the Negev, P.O. Box* 653, *Beer Sheva* 84 120, *Israel.*

APPENDIX 1

The distribution of $D_t(i)$ was derived by

$$D_t(i) = \sum_{\text{all } M} P(M) \binom{M}{i} [P_t(M)]^i [1 - P_t(M)]^{M-i}$$

where $P_t(M)$ is the probability that a person intending to request a seat on the full span does so during period t prior to departure, given that a total of M persons will have requested space by the time of departure.

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 6, 1977. Pergamon Press. Printed in Great Britain

Production Technology and Management in Planned and Market Economies

CC GALLAGHER
The University of Glasgow

(*Received April 1977: in revised form May 1977*)

This paper discusses how technological innovation takes place in manufacturing industry, and especially in the engineering sector, contrasting in particular the effects of the Western market economy, and the East European planned economy. The relative lack of cross-influence between formal scientific research and actual industrial innovation practice in this area is first discussed. Comparative studies of engineering practice in planned and market economies are described which point to the important influence of the economic environment on innovation practice in the factory's design office, and on the shop-floor. The East European Soviet Type Economy is seen to have features which in the light of recent Western empirical studies into industrial innovation, must hamper this process. In particular it is proposed that its relative economic and social rigidity has an important restricting influence on the actual process of innovation in the factory.

1. THE LACK OF RELATIONSHIP BETWEEN SCIENCE AND MANUFACTURING TECHNOLOGY

THE ASSUMPTION has always been made that there is a strong interrelationship between science and the process of industrial production; in practice this does not appear to be valid. It is important at this point to distinguish between the practice of research which is carried out on the fundamental problems of science, principally in places like universities, and the practice of development work and design work, which is carried out principally in manufacturing industry.

A number of recent studies [10, 7, 3], have confirmed that science tends to develop from previous science, and technology from previous technology, and that any cross-linking is of only marginal importance. Of course the extent of any cross-linking will depend to some degree upon the industry type. Thus in the engineering industries it is small, while in pharmaceuticals it will be more important.

Thus in a study of publications [7, p. 97] Price concluded that science builds mostly on earlier science, and technology mostly on earlier technology, and that direct inter-connections between the two were quite rare, and Langrish in his important study of innovation in British firms came to a similar conclusion

[10, p. 1]. "We have paid particular attention to the relation of basic science to innovation. To our minds, our failure to find more than a small handful of direct connections is the more striking for the fact that we set out deliberately to look for them. Our conclusions on this point have proved unpopular in some quarters. Some academic scientists find it difficult to accept what most businessmen already know: that the great bulk of basic science bears only tenuously if at all on the operations of industry". He concludes that the relationships between science and technology are by no means simple or direct, and suggests that the three most important effects of science are:

(a) Curiosity orientated science in academic institutions provides techniques of investigation.
(b) Science enters innovation already embodied in technological form, i.e. it helps technology to build on technology.
(c) It also provides people trained in scientific techniques and ways of thought, and this is probably its most important influence.

It would appear then, that there is little immediate cross-influence between a nation's scientific work and its industrial expertise.

Vol. 5
1977

This conclusion is of particular interest in relation to the Soviet Union and the countries of Eastern Europe in that it may explain the discrepancy which exists between their often high scientific achievement and more doubtful technological performance. Sutton in his recent controversial book about the development of Soviet industry, found this discrepancy to be particularly marked: "The Soviet problem is not that the nation lacks theoretical or research capability or inventive genius. The problem is rather that there is a basic weakness in engineering skills, and the system's mechanisms for generating innovation are almost non-existent .... The major conclusions presented by this study are that Western technology has been, and continues to be the most important factor in Soviet economic development" [13].

Thus Sutton puts forward an 'inability hypothesis' that the spectrum of engineering skills required of a modern manufacturing nation do not exist in the Soviet Union in sufficient quantity. It is suggested here that an alternative explanation is that, while the skills exist or are potential, the economic and social environment does not allow their full exploitation and development.

2. A COMPARISON OF PRODUCTION TECHNOLOGY IN A PLANNED AND A MARKET ECONOMY

Having argued that a nation's industry technology is not influenced dramatically by its research output, it is suggested that what does strongly influence it is the economic and social environment in which it exists. The results of the

comparative studies described below indicate that small but significant technological and managerial differences occur in manufacturing industries which operate in different economic environments.

While some of the effects of these technical and manufacturing differences may be small at the level of the factory, they can be seen as an indication of the ability of the system to absorb change and innovation.

Thus a gear, a shaft, or an engineering assembly designed in the factory of a market economy will exhibit small but significant differences when compared with similar components designed in the factory of a planned economy, and there will be differences in the methods used to make these technical parts.

Although production technology uses the principles of science extensively, the practising design and production engineer must still reach many decisions on the basis of past experience, or the formal and informal rules of the particular social–economic system in which he works. One obvious example is whether the particular national system of engineering standards used in a society is compulsory as in a planned economy or voluntary as in a market economy. A second example would be whether a market-controlled or state-controlled system of pricing existed.

In a study carried out in Birmingham and Warsaw [4] the procedures and practices followed for the design and manufacture of certain engineer's cutting tools was compared for Poland and the UK. The study looked in detail at design features, types and quality of steels used, relevant national engineering standards, production quantities, manufacturing times, and the design recommendations of the technical literature. The Polish engineers were found to have used a 'classical' or 'textbook' approach to their problems. They designed for optimum performance regardless of other considerations. In a similar way the sequences of manufacturing processes were also specified in a 'textbook' or 'academic' manner. The British designers, in contrast, took a less rigid approach to design features in order to simplify and reduce manufacturing costs, and where possible on the basis of experience, used short-cuts to eliminate manufacturing processes and reduce the final cost of the product. To give one example, in the design of a keyway broach the classical academic design of the Poles was seen in its British counterpart to have had non-essential design features and production processes eliminated—possibly at some risk—in an effort to reduce the final cost of the product, and it is a reflection of the skill of the manufacturer that he was able to predict reliably when corners could be cut in such a way. The Polish keyway broaches were found to be more expensive in terms of material, machine-time and labour.

A similar comparative study of the industrial standards systems of Poland and Britain [6] also revealed differences which must have a significant influence on the decisions and motivations of practising production technologists and managers.

Superficially, the industrial standards systems of East and West appear to

serve the same purpose, but on closer examination, fundamental differences are found. Western standards are first of all intended as a strongly recommended guide for manufacturers and users, in order that they may find a minimum of incompatibility between products from different sources. They are secondly intended to act as a guide to what industry currently feels are reasonable quality, accuracy, and safety parameters, in the light of the then existing technology. Finally they are intended to encourage industry to take advantage of the economic benefits of standardisation. They tend to avoid detail and manufacturing technology, except where it is essential to the standard.

East European standards have a quite different and much more important function in the economic system. First because there is no competitive market mechanism, they are almost the only means of control a customer has over the quality of the products he purchases, and for this reason state standards are compulsory. Second, because they often give considerable detail about manufacturing dimensions and tolerances, they partially act as instructions to the factory about what must be made and how, and to a certain extent therefore, replace completely or partially factory design and planning procedures. Finally, because they present a prospective consumer with a fixed range of choices, they act as a purchasing catalogue. They therefore also compel both manufacturer and purchaser towards the economic benefits of standardisation. These contrasting roles are reflected in many of the differences to be found between the two types of standards.

Vol. 5
1977

A legally enforceable system of standards plays a unique role in a planned economy country in that it replaces the market mechanism to ensure that acceptable levels of quality are maintained. Such a system could also be a source of economic inefficiency by demanding excessively high levels of quality. As with most economic and technological techniques, the use of standards and standardisation demands a form of compromise. On one side it can lead to reduced variety and the lower costs associated with mass-production. On the other if used excessively and inflexibly it can restrict good design, hamper creativity and innovation and limit the conditions for rapid change which a modern technological environment needs.

In the Polish planned economy, the three function groups of design, manufacture and user influenced each other much less in the area of reducing the cost of the final product than was the case in the British counterparts. There was no consistent overall external impulse which compelled the Polish functions to collaborate with a unified outlook towards product improvement and therefore each group tended to behave as an isolated unit.

The effects of the planned economy environment on an engineering designer results in a number of significant differences when compared with his market-economy counterpart. He lacks cost consciousness and he designs principally for technical excellence, because there is no effective motive for cost reduction. His range of choice of raw materials and components is considerably more limited,

as there is virtually only one source of supply. But, conversely such limited choice does result in easier variety control.

The overall picture which emerges from these comparisons is that the system rigidity which results from a Soviet Type Economy at the level of the design office and production planning department in an enterprise, inhibits innovation and technical change.

3. INNOVATION AND THE FACTORY ENVIRONMENT

Several detailed studies have been carried out in the West in recent years to try to find out why some nations and firms innovate consistently, while others are less successful. Arrow [1] in an important paper has also considered this problem from a theoretical viewpoint, and demonstrated the essential link between technological progress and modern information theory. "The understanding of the transmission of knowledge is of especial importance in two of the key socio-economic problems of our time: (a) international inequalities in productivity, and (b) the failure of the educational system in reducing income inequality. The two problems have a considerable formal similarity. If one nation or class has the knowledge which enables it to achieve high productivity, why is not the other acquiring that information? That a nation or class has a consistently high productivity implies a successful communication system within the nation or class, so the problem turns on the differential between costs of communication within and between classes (and nations)".

It is relevant to note apparent Western success factors, then look at planned economy industry, to see whether they appear there as strengths or weaknesses. Three factors are most frequently present in circumstances of successful innovation.

First, there is flexibility within an organisation. Langrish describes how he found that in large organisations, small teams of from two to ten people were set up in reaction to a stimulus, from a 'pool' of people on development work. "These *ad hoc* groups tend, however, to overlap in personnel. The 'pool' remains fluid in the sense that the groupings in it undergo continuous restructuring" [10, p. 22].

Second, there is flexibility between organisations by means of the movement of people. This may appear at first to be a trivial point, but in practice has been found to be the most important mechanism of technology transfer. "Of the 158 important ideas made use of in 51 innovations, 102 came from outside the firm, . . . and . . . the most important route for the transfer of ideas was the recruitment of a person into the firm" [10, p. 77]. Indeed, Langrish found such so-called 'on the hoof' transfer by means of the movement of a technologist with his background and expertise to be easily the most important method of technology dissemination, and Arrow reached the same conclusion. "While the mass

media play a major role in alerting individuals to the possibility of an innovation, it seems to be personal contact that is most relevant in leading to its adoption. Thus the diffusion of an innovation becomes a process formally akin to the spread of an infectious disease" [1, p. 33].

The third factor is that the 'top person' in the organisation should be committed to the implementation of the new technology. Such a person is necessary because innovation is not always attractive to an established organisation. There will always be strong pressures which encourage stability and hinder change. "It is because resistances . . . have to be overcome that successful innovation often owes much to individuals who commit themselves to pushing a project through" [10, p. 10].

The degree of flexibility which exists in the planned economy countries with reference to the job mobility of technologists both within and between enterprises is known to be significantly less than in the West [5]. Eastern Block managerial job mobility is very largely based on nepotistic procedures, combined with the more rational manpower planning and selection procedures evolved by the personnel sections in enterprises. This can be compared with the Western semi-nepotistic procedures which run in parallel with a form of 'market' selection. The process of selection is always from above, and the selector holds the initiative with his knowledge of potential candidates.

Vol. 5
1977

There exist in a planned economy industry many rules, regulations and procedures which are common throughout all of that industry. Their importance lies not in their existence, for rules, regulations and procedures performing similar functions can be found in all market economy industries, but in their common or standard basis.

National standardisation has both advantages and weaknesses. When a rule or regulation is good, it is universally good. Unfortunately where such rules are bad, or are only truly relevant to a small sector of industry, or cannot be modified easily as circumstances change, then their effect can be disastrous. Thus the state holiday arrangements in Poland are an example of national regulations working reasonably well, and the state-regulated system of payment in Poland is an example of a regulation creating problems because of lack of flexibility. It appears therefore that there is less work flexibility and job mobility both within and between enterprises in a planned economy, and if Western research findings are relevant, then this must result in a significant restriction in that society's ability to generate technological change in industry.

4. SOME ECONOMIC ASPECTS OF TECHNOLOGICAL CHANGE

The size of Western firms has also been studied in relation to their ability to innovate, and the assumption that large corporations are more successful than

small firms in this field has been shown to be unfounded. "Despite the views of Galbraith, Schumpeter and others, there is little evidence that industrial giants are needed in most industries to ensure rapid technological change.... On the contrary, there seem to be considerable advantages in a diversity of firm sizes ..." [11]. Jewkes also concluded that "... innovation does not flow wholly, or even mainly, from the largest corporations. The very size of a corporation may make it in some ways a centre of resistance to change" [9, p. 13]. Finally Mansfield and his colleagues concluded from a detailed study of the rate of innovation of numerically controlled machine-tools in American industry that: "Judging from our results concerning numerically controlled machine-tools, the rate of diffusion tends to be more rapid in less concentrated industries" [11, p. 224]. In other words those with many small firms rather than a few large corporations. The size structure of Soviet and East European industry has a preponderance of large organisations, and a disproportionately small number of small enterprises by Western standards, and thus on the basis of Western experience would not appear to lend itself to rapid technological change.

It has also been noticed in the West that there is a tendency in large organisations to create tasks for a sector which for some reason has lost its normal work-load, and that the talents of the sector may be grossly mis-matched with the demands of the new work. A striking example occurred in Britain with the run-down of the state nuclear energy programme in the 1960s. Jewkes [9, p. 16] commented that "as the case for nuclear energy well illustrates, when Civil Service permanent establishments become too large for their tasks, the line of least resistance will naturally be that of trying to make work for redundant staffs". The same action to stem the rundown of the British Atomic Energy Authority's work force, was also attacked from a technological point of view by Edwards. "The £670,000 allocated to the Royal Aircraft Establishment for numerical control of machine-tools, and the £300,000 allocated to the Atomic Weapons Research Establishment for group technology (a manufacturing technique) are examples by governments to stimulate the wider acceptance and use of new concepts .... The temptation to channel large amounts of money towards government establishments which have only small amounts of their former work remaining has seemed to be overpowering, so that the real specialists in these fields, many of whom are industrial companies, have been side-stepped, ignored, and in some cases insulted" [2, p. 35]. It is unlikely that the large state organisations of the Soviet Union and Eastern Europe are immune from such temptations.

It has been pointed out by Salter [12], that perhaps the most important type of technological change takes the form of the many small improvements which can be continuously diffusing throughout industry relatively quickly, rather than the one-off major technological break-throughs, which by their nature are more apparent to the outside observer.

Such small innovations can however generate manning problems. Each small

change may not in itself warrant a reduction in employment, but in total over time the accumulated small changes may imply a significant reduction. Unless this is perceived and acted upon, serious overmanning may and frequently does build up [8, p. 283]. Given the relatively inflexible nature of planned economy industry and the ideological temptation to create hidden in preference to overt unemployment, it is unlikely that this causes less of a problem than it does in capitalist countries.

5. CONCLUSIONS

1. The relatively poor industrial manufacturing performance of the Soviet Union in particular, in contrast to its established scientific ability, can be explained in Western terms by the lack of any immediate influence by science on manufacturing industry.
2. For effective innovation and technological change to take place a flexible and fluid social and economic environment is needed within the factory. The rigidity imposed by a Soviet-type economic system does not provide such an environment.
3. The type of economic and social environment in which an industry operates, significantly affects detailed technology in the factory, both in terms of design and manufacture.
4. Sutton's 'inability hypothesis' proposes that the Soviet Union's relatively slow rate of innovation is the result of the simple lack of basic engineering skills. It is suggested here that rather than there being a lack of engineering skills, the very operation of a Soviet Type Economy because of its social and economic rigidity blunts any existing skills and prevents their full expression.

The second and third points listed above pose the question of just how significant these influences are. Salter [12], when discussing the mechanism of technical change, pointed out that ". . . in fact many experienced observers rate the cumulative effect of small unnoticed modifications and improvements as equally great as the more significant changes normally regarded as innovations". If this is a reasonable statement, then the "small unnoticed modifications and improvements" (or lack of them) which can be directly attributed to the effects of the social and economic environment are of great importance to a country's industrial efficiency.

Vol. 5
1977

REFERENCES

1. Arrow JK (1969) Classificatory notes on the production and transmission of technological knowledge. *Am. Econ. Rev.* **59**(2), 29–35.
2. Edwards GAB (1971) *Readings in Group Technology*. Machinery Publishing, London.

3. FORES M (1975) Innovation in the engineering sectors. Royal Economic Society Conference, Government and Innovation, 1975, Cambridge.
4. GALLAGHER CC (1973) How economic systems affect technology. *Chartered Mech. Engr* **20** (7), 69–72.
5. GALLAGHER CC (1973) Communist factory environment. *Works Mgmt* **26**(9), 5–10.
6. GALLAGHER CC (1973) Industrial standards in planned and market economies. *Prodn Engr* **52**(11), 421–425.
7. GRUBER WG and MARQUIS DG (eds.) (1969) *Factors in the Transfer of Technology*, p. 97. M.I.T. Press, Cambridge, MA.
8. HUNTER LC, REID GL, BODDY D (1970) *Labour Problems of Technological Change*. Allen & Unwin, Edinburgh.
9. JEWKES J (1972) Government and high technology. Institute of Economic Affairs, Occasional Paper 37. London.
10. LANGRISH J, GIBBONS M, EVANS WG, JEVONS FR (1972) *Wealth from Knowledge*. Macmillan, London.
11. MANSFIELD E (1968) *The Economics of Technological Change*. Norton, New York.
12. SALTER WEG (1969) *Production and Technical Change*, 2nd edn. Cambridge University Press, U.K.
13. SUTTON AC (1973) *Western Technology and Soviet Economic Development*. Vol. 111, 1954–1965. Hoover Institution Press, Stanford University.

ADDRESS FOR CORRESPONDENCE: *Dr CC Gallagher, Dept of Management Studies, The University of Glasgow, Adam Smith Building, Glasgow G12 8RT, UK.*

BLANK PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 6, 1977. Pergamon Press. Printed in Great Britain

Memoranda

Management Science Techniques in Accounting: An Empirical Study

ACCOUNTING has expanded in scope to cover many areas of analytical techniques. They include sampling, linear programming, regression analysis, and statistical decision theory, goal programming, pert-cost, simulation, and markov chains. This is evident by a variety of reports and publications dealing with the application of operations research in the traditional accounting topics. As early as 1964, a Committee of Measurement in the Institute of Management Sciences suggested:

> The expected development and growth of management services will take place only if the accountant of the future possesses more sophistication in mathematics than he typically has had in the past. Improved methods of analysis now classified as 'operations research' or 'management science', must become the concern of the accountants [1].

The expanded role of analytical techniques in accounting is further evidenced by the Certification in Management Program offered by the Institute of Management Accounting. This program fully recognizes the inclusion of analytical techniques as part of its requirement. Furthermore, many CPA firms have added a quantitatively oriented Management Services Division to further integrate the areas of analytical techniques into the accountants' function.

The purpose of this paper is threefold. The first is to outline the significant areas of analytical techniques as applied to accounting by accountants of Fortune 500 firms. Secondly, an analysis is made of the characteristics of the firm and the extent of the use of analytical techniques by accountants of major United States corporations. Finally, based on the survey, some recommendations are presented.

TABLE 1. RELATIVE APPLICATION OF MANAGEMENT SCIENCE TECHNIQUES

Management science techniques	Mean
Linear programming	3·10
Goal programming	1·99
Other mathematical programming techniques (dynamic, nonlinear, integer)	2·18
Learning curves	2·53
Simple linear regression	4·16
Multiple linear regression	3·78
Nonlinear (simple & multiple) regression	3·00
Statistical sampling	3·56
Statistical forecasting (e.g. time series analysis)	4·14
Statistical variance analysis	3·24
Decision theory (e.g. expected monetary value)	3·05
Net work analysis (e.g. Pert/CPM)	3·05
Inventory models	3·46
Queuing theory	1·96
Markov analysis	1·46
Computer simulation models	4·17
Computerized ratio analysis	3·44
Grand mean	3·15

TABLE 2. SIZE OF ORGANIZATION AND THE USE OF MANAGEMENT SCIENCE TECHNIQUES

Management science techniques	Employees (in thousands)				Gross sales (in millions)						Use of consultants		
	0–49	50–99	100–199	Over 600	$0–$199	$200–$599	$600–$799	$800–$999	$1000–$5999	Over $6000	Yes	Unsure	No
	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Linear programming	2·81	2·63	3·50	4·05	1·00	2·67	2·15	3·33	3·00	4·43	3·58	2·14	2·52
Goal programming	1·98	1·93	1·80	2·22	1·00	2·67	2·27	1·58	1·84	2·64	2·07	2·00	1·91
Other mathematical programming	2·07	2·18	2·25	2·42	1·00	2·67	1·36	2·45	2·14	2·79	2·64	1·57	1·62
Learning curves	2·30	2·93	3·60	2·50	2·00	2·67	1·77	1·60	2·86	2·50	2·54	3·43	2·41
Simple linear regression	3·98	4·65	4·83	4·00	4·00	4·00	2·54	3·58	4·53	4·57	4·58	3·86	3·53
Multiple linear regression	3·46	4·18	4·83	3·85	2·00	3·67	2·15	3·18	4·07	4·71	4·37	3·29	2·94
Nonlinear regression	2·68	2·88	3·67	3·60	1·00	4·67	1·69	2·45	3·10	4·00	3·56	1·71	2·29
Statistical sampling	3·19	3·71	4·00	4·00	4·00	4·67	2·38	2·83	3·72	4·36	3·90	2·57	3·06
Statistical forecasting	3·68	4·88	4·17	4·50	5·00	4·33	3·15	3·42	4·37	4·57	4·60	3·29	3·47
Statistical variance analysis	2·77	4·06	3·67	3·55	1·00	3·67	3·08	2·83	3·25	3·79	3·55	1·86	2·94
Decision theory	2·56	3·88	3·83	3·25	1·00	3·00	2·08	2·25	3·33	3·57	3·40	3·43	2·35
Network analysis	2·71	3·12	4·50	3·25	4·00	3·33	2·38	2·75	3·03	3·86	3·42	3·67	2·38
Inventory models	3·12	3·59	5·50	3·58	2·00	2·67	2·62	3·42	3·58	4·07	4·03	3·29	2·47
Queuing theory	1·77	1·88	3·20	2·05	1·00	2·33	1·38	1·50	1·96	2·86	2·26	1·86	1·53
Markov analysis	1·32	1·25	1·60	1·84	1·00	2·00	1·38	1·00	1·48	1·79	1·61	1·43	1·18
Computer simulation models	3·74	4·71	6·17	4·05	1·00	4·67	2·15	4·08	4·35	5·50	4·63	3·86	3·44
Computerized ration analysis	3·05	4·23	3·83	3·68		5·00	2·54	2·82	3·75	3·36	3·67	2·60	3·18
Grand mean	2·78	3·34	3·91	3·33	2·00	3·45	2·19	2·67	3·35	3·73	3·46	2·69	2·69
Sample size	55	25	10	20	6	7	12	11	60	14	62	13	35

MANAGEMENT SCIENCE TECHNIQUES IN ACCOUNTING

The information concerning the type of management science techniques and the extent of the use of these techniques by accountants was collected by means of a questionnaire. A questionnaire was mailed to 340 large Fortune 500 firms and 110 firms responded to the questionnaire. The firms were asked to rank the extent of the application of management science techniques on a seven point scale, in which 1 represented no use of the technique and 7 an extensive use of the technique. Table 1 shows the extent of the use of management science techniques by the firms.

SIZE OF THE ORGANIZATION AND THE USE OF MANAGEMENT SCIENCE TECHNIQUES

Table 2 shows the characteristics of the organization and the extent of the application of management science techniques in major United States Corporations. The application of quantitative techniques in budgeting and profit planning goes up in the beginning as the gross sales of the organization increases, decreases in the organizations having sales of $600–800 million dollars and then the application of the techniques starts increasing with the increase in gross sales. As the number of the employees in an organization goes up, the application of quantitative techniques in profit planning increased with the exception of the organizations employing over 600,000 employees. Nevertheless, the use of sales forecasting techniques in profit planning is used more widely in the largest organizations in terms of number of employees. Table 2 also shows that the organizations using consultants periodically make more use of analytical techniques than the organizations which do not use or are unaware of consultants in the organization.

MANAGEMENT SCIENCE TECHNIQUES IN DIFFERENT INDUSTRIES

The firms engaged in producing goods and/or services in the areas of leisure time, container publication, personal care products, and drugs make the greatest use of management science techniques in budgeting and profit planning. Whereas the firms engaged in producing telecommunication, general machinery, steel, and tire and rubber apply the quantitative techniques the least in profit planning. Table 3 illustrates the extent of the use of each specific technique in a specific industry and the order in which management science techniques are being applied by various industries.

CONCLUSION

As stated previously that the computer simulation models, forecasting, statistical sampling, statistical variance analysis, simple and multiple regression, linear programming, and computerized ratio analysis are being used widely in profit planning by major US corporations. Hence, it is recommended that accounting professors involved in this area emphasize the application of at least these techniques to accounting.

REFERENCE

1. IMS Committee comments for Common Body of Knowledge Study. *Jl Accountancy* (December, 1964), pp. 79–83.

Lal C Jagetia
(*Received May* 1977)
The James J Nance College of Business Administration
Cleveland State University
Cleveland
Ohio 44115
USA

Table 3 of this paper is on pp. 744 and 745.

TABLE 3. EXTENT OF THE APPLICATION OF MANAGE

Management science techniques	Aerospace	Appliances	Automotive	Beverages	Building materials	Chemicals	Conglomerate	Containers	Drugs	Electronics	Food	General machinery
	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
Linear programming	2·00	2·40	2·17	1·67	3·56	4·25	2·44	4·00	3·11	2·22	2·62	1·33
Goal programming	2·67	2·80	2·91	4·50	1·22	1·45	2·63	2·25	3·25	2·43	2·62	2·00
Other mathematical programming	1·50	2·60	2·25	1·67	2·13	2·23	2·44	2·80	2·33	1·43	2·46	1·33
Learning curves	3·80	3·00	3·25	2·67	2·13	2·50	2·63	2·67	2·75	3·11	1·83	2·00
Simple linear regression	3·40	4·00	4·25	3·33	5·11	4·54	4·63	5·40	4·22	3·22	3·23	2·83
Multiple linear regression	3·40	3·80	4·33	3·33	4·67	3·92	4·06	5·00	3·50	3·00	3·23	2·50
Nonlinear regression	2·60	3·20	2·92	2·33	3·11	3·54	3·44	4·00	4·13	2·00	2·85	2·00
Statistical sampling	3·20	3·40	3·92	2·67	2·11	3·62	3·69	4·20	3·89	3·11	4·23	2·67
Statistical forecasting	3·60	4·20	4·42	4·33	5·11	4·15	4·69	5·20	5·00	3·00	4·08	2·83
Statistical variance analysis	3·40	4·40	3·67	4·67	3·22	3·15	3·94	5·00	3·78	2·56	3·54	2·67
Decision theory	4·40	3·00	3·83	3·33	2·89	2·54	3·50	4·00	3·33	2·78	3·31	3·00
Network analysis	3·40	4·60	3·55	4·33	2·78	3·08	3·25	4·40	3·89	3·89	2·54	2·67
Inventory models	4·50	4·40	3·45	4·00	3·89	3·23	3·38	4·40	4·11	4·00	3·31	2·67
Queuing theory	2·33	3·20	2·27	1·50	1·89	1·85	2·06	2·75	2·00	2·25	1·85	2·00
Markov analysis	1·00	1·40	1·45	1·00	1·44	1·54	1·50	2·20	1·38	1·63	1·33	1·33
Computersimulation models	4·80	4·00	4·08	2·67	5·11	4·23	4·38	4·40	4·44	3·33	3·38	3·17
Computeized ratio analysis	4·50	3·40	4·00	4·00	3·33	3·64	4·56	3·50	4·67	3·13	3·64	2·67
Grand mean	3·29	3·40	3·35	3·09	3·17	3·15	3·42	3·96	3·53	2·80	2·95	2·33
Sample size	5	5	12	3	9	13	16	5	9	9	14	6

Vol. 5
1977

MENT SCIENCE TECHNIQUES BY DIFFERENT INDUSTRIES

Instruments	Leisure time	Metals and mining	Miscellaneous manufacturing	Energy	Office equipment	Personal care products	Publishing	Services engineering	Special machinery	Steel	Textile	Tire and rubber	Telecommunication	Other
Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean	Mean
1·75	3·00	2·78	2·75	4·38	2·67	2·83	2·33	3·00	2·38	2·00	2·67	6·00	1·33	3·36
4·75	4·33	2·71	2·57	1·62	2·67	3·20	6·50	2·50	1·33	2·00	1·00	2·00	4·67	2·00
1·75	3·33	2·22	2·75	2·38	1·50	2·30	2·00	2·00	2·00	1·20	2·00	2·00	1·67	2·08
2·75	3·67	3·00	2·43	1·46	4·17	1·75	3·00	2·25	3·00	2·00	2·33	1·00	3·00	1·85
3·00	5·33	4·67	4·50	3·69	4·00	3·67	4·00	4·00	3·63	4·00	4·00	5·00	2·00	3·64
2·75	4·67	4·67	4·38	3·23	3·33	3·67	3·67	4·00	3·75	3·80	3·00	2·00	1·67	3·36
1·50	4·67	3·56	3·63	3·15	2·33	3·17	3·00	3·00	2·50	2·60	3·00	3·00	1·67	2·29
3·25	5·33	3·44	4·50	3·69	3·67	5·00	4·00	3·80	3·63	2·20	4·33	2·00	2·67	3·29
4·00	6·00	4·78	5·38	3·15	4·50	5·67	5·00	3·40	4·00	3·60	5·33	2·00	3·33	4·07
3·25	5·67	3·89	4·50	2·77	4·00	4·83	4·67	3·00	3·13	2·40	3·67	2·00	3·33	3·29
4·00	4·33	3·67	3·63	2·77	3·33	3·00	4·67	3·40	2·88	2·40	3·33	2·00	3·33	3·00
4·50	5·00	3·78	4·25	2·77	4·00	3·83	4·33	3·20	3·38	2·80	3·00	1·00	4·00	2·57
4·25	5·33	3·75	4·43	3·00	4·67	4·50	4·67	2·60	3·50	2·60	3·33	2·00	5·00	3·46
2·75	4·00	2·25	3·14	2·23	2·00	2·50	2·67	2·20	1·86	1·40	1·67	3·00	2·67	1·77
1·50	2·67	1·88	1·50	1·62	1·33	1·60	1·67	1·80	1·14	1·00	1·67	1·00	1·33	1·08
3·75	5·33	3·89	4·88	5·00	4·50	4·33	3·67	4·00	3·38	4·40	2·33	5·00	4·33	3·93
5·00	5·33	3·22	4·83	2·58	4·80	3·60	3·67	3·40	3·13	3·60	2·50	3·00	4·50	3·43
3·21	4·59	3·45	3·80	2·91	3·13	3·57	3·69	3·05	2·91	2·58	2·96	2·59	2·14	2·87
4	3	9	8	13	6	6	3	5	8	5	3	1	3	14

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 6, 1977. Pergamon Press. Printed in Great Britain

The Sensitivity of the Backorder Inventory Model to Lot Size Errors

IT IS a well known fact that for the lot size inventory system under certainty with infinite backorder costs, the annual cost, $K(Q)$, of holding inventory and ordering lots of size Q is relatively insensitive to the value of Q. For example, if yearly requirements are ordered in lots which are 25% larger than the optimal order quantity, Q^*, then annual costs, $K(1{\cdot}25Q^*)$, will be only 0·025 higher than $K(Q^*)$. In general,

$$\frac{K(Q)}{K(Q^*)} = \frac{1}{2}\left[\frac{Q}{Q^*} + \frac{Q^*}{Q}\right] \qquad (1)$$

gives the ratio of annual cost of the sub-optimal quantity to the annual cost of ordering the optimal quantity. Of course, when backorder costs are not infinite, it will be optimal to incur some shortages near the end of the cycle.

In this paper a formula comparable to (1) will be developed for a lost size system when the cost of backlogging is finite. The behavior of the inventory system is illustrated in Fig. 1 and is taken directly from Hadley and Whitin [1].

Demand is at a constant rate of R units per year; t_1/T represents the fraction of a cycle over which inventory is positive, while t_2/T represents the fraction of a cycle over which demand is, essentially, being postponed until the next delivery arrives. By the time the replenishment order does arrive, the number of units in the backlog is equal to s.

If h is the cost of holding a unit in inventory for one year, then

$$C_h(Q, s) = \frac{h(Q-s)^2}{2Q} \qquad (2)$$

is the annual cost of holding average inventory. Similarly, if b is the cost of backlogging a unit of demand for 1 year, then

$$C_b(Q, s) = \frac{bs^2}{2Q} \qquad (3)$$

is the annual cost of backlogging, where (3) is based on the average number of units backordered per cycle. If the cost of placing an order is equal to A, then the annual cost of ordering is

$$C_0(Q, s) = \frac{AR}{Q}. \qquad (4)$$

Hence, the total cost of holding inventory, backlogging demand and ordering for 1 year is

$$K(Q, s) = C_h(Q, s) + C_b(Q, s) + C_0(Q, s). \quad (5)$$

Using well known methods, the optimal values of Q and s are given by

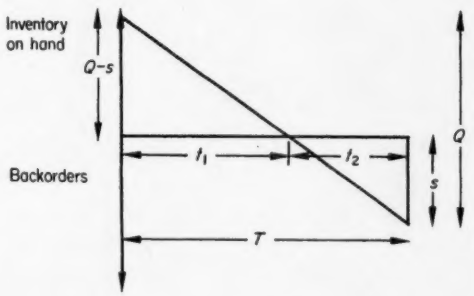


FIG. 1.

Vol. 5
1977

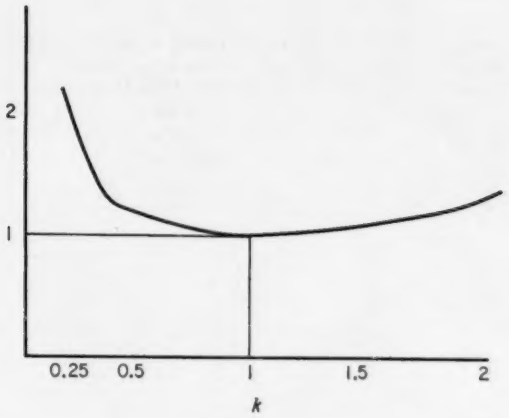


FIG. 2.

$$s^* = \sqrt{\frac{2RA}{b}}\sqrt{\frac{h}{h+b}} \tag{6}$$

and

$$Q^* = \sqrt{\frac{2RA}{h}}\sqrt{\frac{h+b}{b}}. \tag{7}$$

Hence, the optimal on hand inventory after replenishment is

$$(Q^* - s^*) = \sqrt{\frac{2RA}{h}}\sqrt{\frac{b}{h+b}}. \tag{8}$$

In order to find $K(Q, s)/K(Q^*, s^*)$, an equation is needed for $K(Q^*, s^*)$. The algebra required to find this ratio can be simplified if it is recognized that, at optimality,

$$\frac{h(Q^* - s^*)^2}{2Q^*} + \frac{bs^{*2}}{2Q^*} = \frac{RA}{Q^*}. \tag{9}$$

Hence, at optimality, $K(Q^*, s^*) = \dfrac{2RA}{Q^*}$, or

$$K(Q^*, s^*) = \sqrt{2RAh}\sqrt{\frac{b}{h+b}}. \tag{10}$$

Now, dividing equation (5) by (10) yields

$$\frac{K(Q, s)}{K(Q^*, s^*)} = \frac{h(Q-s)^2}{2Q}\sqrt{\frac{h+b}{b}}\sqrt{\frac{1}{2RAh}} + \frac{bs^2}{2Q}\sqrt{\frac{h+b}{b}}\sqrt{\frac{1}{2RAh}} + \frac{RA}{Q}\sqrt{\frac{h+b}{b}}\sqrt{\frac{1}{2RAh}}$$

After some algebra and using the results in (6), (7) and (8), equation (11) simplifies to

$$\frac{K(Q, s)}{K(Q^*, s^*)} = \frac{1}{Q^* - s^*}\left[\frac{Q}{2} - s + \frac{s^2}{2Q}\right] + \frac{1}{2Q}\left[\frac{s^2}{s^*} + Q^*\right]. \tag{12}$$

Equation (12) is a function entirely of Q^*, s^*, Q and s. If backorders are prohibited, then it can be shown that (12) and (1) are identical. Equation (12) is plotted for values of k which satisfy $Q = kQ^*$ and $s = ks^*$ in Fig. 2. As expected, the curve is flat over a wide range of values for k. In fact, the ordinates of $K(Q, s)/K(Q^*, s^*)$ are the *same* as those of $K(Q)/K(Q^*)$ (for the lot size model without backorders), if $k = Q/Q^*$. That is, if Q and s are not the optimal values of Q^* and s^*, then the percentage by which actual cost exceeds minimum cost for the year is exactly equal to that which prevails for the simple lot size system when $k = Q/Q^*$.

REFERENCE

1. HADLEY G and WHITIN TM (1963) *Analysis of Inventory Systems*, p. 43. Prentice-Hall, New York.

Robert M Brown
(*May* 1977)

University of Wisconsin
Commerce Building
1155 *Observatory Drive*
Madison
Wisconsin 53706
USA

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5, No. 6, 1977. Pergamon Press. Printed in Great Britain

Toward a Theory of Hypergames

THE PURPOSE of this note is briefly to examine an extension to the analytical framework of Game Theory, enabling one to model less 'idealised' situations.

Game Theory is applied on the assumption that the players are well-informed as to the game being played, i.e. of each other's preferences and range of strategies. Though their *interests* may diverge, they thus have a *common perception* of the problem in hand. But in real life—and even in some 'laboratory games'—it is clear that decision-makers' perceptions of the situation may differ radically. In effect, the 'players' are trying to play different games, sometimes with unfortunate results. One cannot satisfactorily model such an interaction using, in the normal way, either Game Theory or the more recent approach of Metagame Theory (though the latter has provided interesting new insights [1, 3], players' decisions are analysed using metagames based on a common game supposedly perceived by all players). We propose a form of analysis based on the notion of a *hypergame*. Roughly, this is a system comprised of a set of games, with each game interpreted as expressing a particular player's idea of what is happening.

First, for comparison, we give formal definitions of a Game and a Hypergame, using preference orderings for the players, for the sake of generality, rather than the utility scales of the original Von Neumann and Morgenstern treatment [5].

An n-*person Game* in normal form is a system consisting of:

G.1: a set P_n, of n elements, interpreted as the *players* of the game,

G.2: for each $p \in P_n$, a non-empty finite set S_p, interpreted as the set of *strategies* available to p,

G.3: for each $p \in P_n$, an ordering relationship O_p, defined over the product space $S_1 x \ldots x S_n$. This is interpreted as *the ordering of player* p's *preferences* over the set of possible outcomes of the game. We may assume partial or complete ordering.

Each player tries to obtain an outcome as highly-preferred as possible, choosing a strategy in the full knowledge of the game, though unaware of the specific choices of the other players.

An n-*person Hypergame* in normal form is a system consisting of:

H.1: a set P_n, of n elements (the players of the hypergame),

H.2: for each $p, q \in P_n$, a non empty finite set S_p^q,

H.3: for each $p, q \in P_n$, an ordering relationship O_p^q defined over the product space $S_1^q x \ldots x S_n^q$.

S_p^q and O_p^q are interpreted respectively as *the set of strategies available for* p, *and* p's *preference ordering, as perceived by* q. That is, they express q's view of p's options and aims. The sets $S_1^q, S_2^q \ldots, S_n^q$ make up q's strategy matrix, and with the set P_n and the orderings $P_1^q, \ldots, O_n^q$ comprise q's *game* within the hypergame, denoted G^q. The hypergame can thus be considered as the set of n games $G^1, \ldots, G^n$.

We define an *outcome of the hypergame* as a strategy n-tuple $s_1^1, s_2^2, \ldots, s_n^n$ where in general s_p^p denotes an element of the set S_p^p. That is, each player chooses from his self-assigned strategy set. This may or may not appear as an outcome of a given player's game. We can thus see that hypergame players —unlike strict game players—may be subject to *strategic surprise*: another player may have more strategies than one had bargained for, so that a completely unexpected result obtains.

As before, each player tries to obtain the most highly-preferred outcome he can, picking a strategy appearing in his own game, with full knowledge of this game. Regarding information of the rest of the hypergame, various possibilities arise. A player may realise that others' interpretations of the situation may differ from his own: he may then have more or less idea of what games the other players are trying to play. Once the information conditions are specified, it is a simple matter to extend the notion of *stability* to hypergames. Roughly, an outcome is stable

when each player perceives that no advantage would be gained by changing his strategy choice. We can apply the usual game-theoretic (or metagame-theoretic) criteria for each player applied to that player's game (or metagames based on it). Since there is not really space to discuss this here, we instead illustrate the general idea of hypergame analysis by means of an example. For simplicity, we choose a 2-person case in which each player has knowledge *only* of his own game.

Suppose our 'players' (p, q) are the rulers of two nations, each desiring peace, but suspicious of the other. Such a situation forms the basis of a well-known cautionary tale (c.f. the 'Andersland'—'Jedesland' story [4]). A search for 'security' prompts an arms race, and the story ends with a "mad rush on both sides to the launching platforms" as each side tries to pre-empt the other's expected attack. Now, we can give a hypergame model of this situation that is a rather compelling alternative to the usual game-theoretic ones. Suppose that, at any given stage in the story, each decision-maker has a straight choice between a co-operative strategy (C) and an aggressive one (A). Both perceive this to be the case: their perceptual games have the same strategy matrices. (This restricted form of hypergame is similar to the 'm-game' idea discussed briefly in [2]). Player p, we suppose, places the four possible outcomes in the following order of decreasing preference:

$\langle C, C\rangle$ 'Peaceful co-existence',
$\langle A, C\rangle$ 'Attack without q retaliating',
$\langle A, A\rangle$ 'Mutual aggression',
$\langle C, A\rangle$ 'Suffer attack by q without replying'.

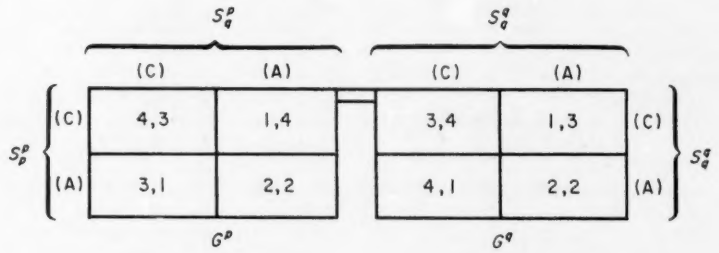
Alas, these preferences are not correctly perceived by q. q imagines p to have the following preference order:

$\langle A, C\rangle$ 'Attack and get away with it',
$\langle C, C\rangle$
$\langle A, A\rangle$
$\langle C, A\rangle$.

Similarly, with the roles of the players reversed. Each really wants only to counter the imagined aggressive tendencies of the other. This 2-person hypergame can be represented by the double matrix shown below.

Looking at the situation from p's point of view, i.e. at the game G^p, it will be seen that q has a dominant strategy—A. So, p may reason, he must act on the assumption that q will adopt this aggressive strategy. He, p, is faced then with a choice between outcomes $\langle C, A\rangle$ and $\langle A, A\rangle$. Of these, the latter is more preferred. So it seems only rational to be aggressive also. The argument for q, looking at G^q is quite analogous.

The important point here is that we can analyse, in a systematic way, the effects—here crucial—of the players' differing perceptions. This cannot be done using a single game to describe what is happening. In the game that would be being played if the players had not mistaken each other's preferences, neither would be tempted to deviate from co-operation. The game that would be being played if each player's preferences *really* were ordered in the more aggressive way assumed above by the other turns out to be the well-known 'Prisoners' Dilemma'. This game—usually used to model this problem—does give a 'rationale' for the 'dash to the launching-platforms'. But then



(Figures represent ordinal preferences, with more preferred outcomes assigned higher numbers)

Hypergame showing perceptions of hostility.

Vol. 5
1977

the essential element of the story—misunderstanding—is left out. Our players' preferences are *not* ordered as in the 'PD' game, *but each thinks the other's are*. Each 'country' is trying only to defend itself against an imagined threat, and would like nothing better than peaceful co-existence. To what extent this turns out to be an accurate representation of any real-world conflict, is, of course, another question.

REFERENCES

1. HOWARD N (1971) *Paradoxes of Rationality*. MIT Press, Cambridge, MA.
2. LUCE RD and ADAMS EW (1956) The determination of subjective characteristic functions in games with misperceived payoff functions. *Econometrica* **24**, 158–171.
3. RADFORD KJ (1977) *Complex Decision Problems*. Reston, Virginia.
4. RAPOPORT A (1960) *Fights, Games and Debates*. University of Michigan Press.
5. VON NEUMANN J and MORGENSTERN O (1944) *Theory of Games and Economic Behaviour*. Princeton University Press.

PG Bennett
(*March* 1977)

S.R.C. Research Fellow
The University of Sussex
Operational Research
Mantell Building
Falmer
Brighton BN1 9RF
UK

BLANK
PAGE

Vol. 5
1977

OMEGA, The Int. Jl of Mgmt Sci., Vol. 5. No. 6, 1977. Pergamon Press. Printed in Great Britain

International Register of Theses and Dissertations

The following are obtainable at the General Library, University of California, Berkeley, CA 94720, USA:

BRETTON GE (1976) The Relationships among Organizational Subunit Size and Selected Aspects of Organizational Communication. PhD Thesis, University of California, Berkeley.

CHEN ET-Y (1976) Capital Asset Prices under Uncertain Inflation. PhD Thesis, University of California, Berkeley.

ELIAS DPB (1975) A Study of Regional Unemployment Differences in Great Britain (1962–1973). PhD Thesis, University of California, Berkeley.

GILL S (1976) Experiments in the Economic Theory of Teams. PhD Thesis, University of California, Berkeley.

GRAUER RR (1975) Risk Aversion and the Structure of Asset Prices: An Examination of Linear Risk Tolerance Capital Asset Pricing Models. PhD Thesis, University of California, Berkeley.

GRAUL PR (1975) Interaction of External Reporting System and Management Decisions. PhD Thesis, University of California, Berkeley.

JAEGER F (1976) Consolidation Strategy for International Air Freight Forwarders: Minimum Cost Routing Problem in a Directed Multi-Commodity Network. PhD Thesis, University of California, Berkeley.

KOWALSZYK SJ (1976) Organizational Structure and Social Responsibility. PhD Thesis, University of California, Berkeley.

LEROY GP (1975) Multinational Corporate Strategy: A Framework of Analysis of Worldwide Diffusion of Products. PhD Thesis, University of California, Berkeley.

MILLOTT JF (1976) Decision Strategies in Hospital Administration: A Management Game Approach. PhD Thesis, University of California, Berkeley.

NG DSP (1975) Essays on the Effects of Information in Financial Markets. PhD Thesis, University of California, Berkeley.

O'REILLY III, CA (1976) The Intentional Distortion of Information in Organizational Communication: A Laboratory and Field Approach. PhD Thesis, University of California, Berkeley.

PERRET FL (1975) Radioactive Waste Storage and Disposal: Methodologies for Impact Assessment. PhD Thesis, University of California, Berkeley.

TYEBJEE TT (1976) Choice Time and Conflict in Consumer Brand Choice Behavior. PhD Thesis, University of California, Berkeley.

WENSTOP FE (1975) Application of Linguistic Variables in the Analysis of Organizations. PhD Thesis, University of California, Berkeley.

WEYANT JP (1976) A Dynamic Model of the United States Energy System: A Tool for Energy and R&D Planning. PhD Thesis, University of California, Berkeley.

WRIGHT FW (1975) Cognitive Information Processing and Investment Analysis: A Descriptive Study. PhD Thesis, University of California, Berkeley.

The following are obtainable at the Joseph Regenstein Library, University of Chicago, 1100 East 57th Street, IL 60637, USA:

BEVERIDGE HS (1975) An Evaluation of the St Louis Model. PhD Thesis, University of Chicago.

BROWN SJ (1976) Optimal Portfolio Choice under Uncertainty: A Bayesian Approach. PhD Thesis, University of Chicago.

DAFT RL (1974) The Process of Organizational Innovation: An Empirical Study of Thirteen High School Districts. PhD Thesis, University of Chicago.

ELLERT JC (1975) Antitrust Enforcement and the Behavior of Stock Prices. PhD Thesis, University of Chicago.

HIRST IRC (1976) Efficiency in Trading and the Revelation of Preferences. PhD Thesis, University of Chicago.

JENSEN KR (1974) An Analysis of Various Rules for Determining Expenditures from Endowment Funds. PhD Thesis, University of Chicago.

MACBETH JC (1975) Tests of the Two Parameter Model of Capital Market Equilibrium. PhD Thesis, University of Chicago.

MEGINNISS JR (1976) Alternatives to the Expected Utility Rule. PhD Thesis, University of Chicago.

MENSCHING JR (1976) Forward Algorithms and Planning Horizons for Deterministic Inventory Problems. PhD Thesis, University of Chicago.

NOLLEN SD (1974) The Supply and Demand for College Educated Labor. PhD Thesis, University of Chicago.

PARSONS BT (1974) Behavior of Prices on the London Stock Market in the Early 18th Century. PhD Thesis, University of Chicago.

RANSON RD (1974) Money, Capital, and the Stochastic Nature of Business Fluctuations. PhD Thesis, University of Chicago.

SALIPANTE PF JR (1975) Organizational Correlates of Disadvantaged Worker Retention. PhD Thesis, University of Chicago.

SCHWERT GW III (1975) Public Regulation of National Securities Exchanges: A Test of the Capture Hypothesis. PhD Thesis, University of Chicago.

VANDAELE WH (1975) The Economics of Crime: An Econometric Investigation of Auto Theft in the United States. PhD Thesis, University of Chicago.

WARNER JB (1976) Bankruptcy Costs, Absolute Priority, and the Pricing of Risky Debt Claims. PhD Thesis, University of Chicago.

The following are obtainable at the Library, Cranfield Institute of Technology, Cranfield, Bedford MK43 0AL, UK:

CHAPMAN M (1971) The Creation of Machine Tool Production Cells by Flow Analysis within an Aerospace Company. MSc Thesis, Cranfield Institute of Technology.

HARFORD DV (1970) The Application of In-Process Gauging to Electrochemical Machining. MSc Thesis, Cranfield Institute of Technology.

HODDINOTT DJC (1972) A Framework for the Development of a Personnel Policy in Life Insurance Company. MBA Thesis, Cranfield Institute of Technology.

Vol. 5
1977

HUQUE KM (1970) Sales Forecasting Procedures for Short-Term Prediction. MSc Thesis, Cranfield Institute of Technology.

JAWAD NS (1971) An Investigation into the Potential Utilisation and Design of a Pick and Place Unit. MSc Thesis, Cranfield Institute of Technology.

JESSUP N (1972) A Preliminary Analysis of the Economics and Alternative Strategies open to British Lion Films Ltd, with Special Regard to Film Distribution in the UK. MBA Thesis, Cranfield Institute of Technology.

JONES FE (1970) A Fluidic Linear Transducer and its Application. MSc Thesis, Cranfield Institute of Technology.

KEE YW (1972) A Feasibility Study for a Food Processing Company in Malaysia. MBA Thesis, Cranfield Institute of Technology.

KEIGHLEY J (1971) An Investigation into Major Factors Affecting Industrial Harmony in a Flow Line Assembly Situation. MSc Thesis, Cranfield Institute of Technology.

KING JDM (1972) Assessment of the Current Position and the Future Prospects of a Small Garage. MBA Thesis, Cranfield Institute of Technology.

SAMPATH K (1971) An Investigation into the Broaching of R.ex 448e and nimonic 80a Super Alloys. MSc Thesis, Cranfield Institute of Technology.

SHERWOOD M (1970) An Appraisal of Group Technology and a Feasibility Study for Implementing the Technique in a Batch-Production Factory. MSc Thesis, Cranfield Institute of Technology.

SINHA RK (1971) Assembly Line Analysis and Shop Loading in an Expanding Company. MSc Thesis, Cranfield Institute of Technology.

SRINIVASARAGHAVAN S (1971) Process Capability Studies and the Relationship Between Capability and Workpiece Parameters. MSc Thesis, Cranfield Institute of Technology.

STANBRIDGE R (1907) The Macro Capability of APT-Based Languages with Particular Reference to Group Technology. MSc Thesis, Cranfield Institute of Technology.

STEPHENSON K (1970) A Study to Plan an Advanced Machining Complex for Manufacturing the Large Component of a Turbogenerator. MSc Thesis, Cranfield Institute of Technology.

SUD PK (1970) Production Control Practice in a Small Firm. MSc Thesis, Cranfield Institute of Technology.

The following are obtainable at the Lyon Playfair Library, Imperial College of Science and Technology, Exhibition Road, London SW7 2BX, UK:

BENTON V (1975) A Financial Model of a Leasing Company. MSc Thesis, Imperial College, London.

BLACKBURNE SL (1975) A Study of Generalised Information System. MSc Thesis, Imperial College, London.

BOUAMRENE MA (1975) Models for Airline Fleet Planning. MSc Thesis, Imperial College, London.

BRIGSTOCKE DHC (1975) A Linear Programming Model for a Decentralised Financial Organisation. MSc Thesis, Imperial College, London.

BUMPSTEED CE (1975) Optimisation of UK Corporation Tax Payments. MSc Thesis, Imperial College, London.

CHAU PKP (1975) Possibilities of Recycled Kraft Paper in the Consumer Market. MSc Thesis, Imperial College, London.

ESPINOSA-RUEDA R (1975) A Non-Linear Approach to the Modelling of Flows of Funds, MSc Thesis, Imperial College, London.

FEI HTH (1975) A Quantitative Review of Job–Shop Scheduling (1967–1974). MSc Thesis, Imperial College, London.

HILTON JMT (1975) The Role of Simulation as a Basis for Defence Decision-Making with an Illustrative Example of a Warship Model. MSc Thesis, Imperial College, London.

HOWARTH MA (1975) GRAFT—An Interactive Graphical Simulation Program Generator. MSc Thesis, Imperial College, London.

JACKSON NW (1975) Medium Term Forecasting of Electricity Demand by Time Series Analysis. MSc Thesis, Imperial College, London.

MILLS JB (1975) The Economic Level of Thermal Insulation in Dwellings. MSc Thesis, Imperial College, London.

MITRA SK (1975) A Financial Simulation Model for a Manufacturing Company. MSc Thesis, Imperial College, London.

PARRY TJ (1975) A Study of some Aspects of International Tax Planning Problems with View to Applying Graph Theory to their Solution. MSc Thesis, Imperial College, London.

POTTER C (1975) The Optimal Pattern of Capacity Increments for Aluminium Smelters. MSc Thesis, Imperial College, London.

RAJENDRAN AN (1975) A Corporate Financial Model of a Canning Firm. MSc Thesis, Imperial College, London.

ROSSI AJ (1975) Public Planning for Large Mining Developments. MSc Thesis, Imperial College, London.

SCOTT SJ (1975) A Study of the Comparative Crime Costs of Cash and Non-Cash Methods of Payment. MSc Thesis, Imperial College, London.

SIKORA ZM (1975) A Study of Vehicle Composition in Concrete Distribution. MSc Thesis, Imperial College, London.

SPREI MA (1975) A Systems Dynamics Model of the Paper Industry. MSc Thesis, Imperial College, London.

SUM PING SK (1975) Reporting System at Cadbury Schweppes Ltd. MSc Thesis, Imperial College, London.

TAN BW (1975) An LP Model of the Multi-Period and Multi-Commodity Blending Problem. MSc Thesis, Imperial College, London.

TORRES-FERNANDEZ HJ (1975) Managerial Reward Systems. MSc Thesis, Imperial College, London.

TSOI SH (1975) A Graph Theoretic Approach to the International Cash Management Problem. MSc Thesis, Imperial College, London.

VARNAVA K (1975) The Effect of Changes in the Delivery Service Level. MSc Thesis, Imperial College, London.

WARD SC (1975) Approximate and Exact Solutions to a Class of Fixed Charge Problems. MSc Thesis, Imperial College, London.

WHELAN VP (1975) Explanatory Models of Aggregate Profit Movements. MSc Thesis, Imperial College, London.

WILKINSON D (1975) Productivity Measurement at Unit Level: A Case Study. MSc Thesis, Imperial College, London.

WONG SH (1975) Aspects of Financial Planning in a Holding Company. MSc Thesis, Imperial College, London.

YANG YP (1975) Forecasting of Bank Clearing Statistics. MSc Thesis, Imperial College, London.

Vol. 5
1977

The following are obtainable at the MIT Microreproduction Laboratory, Massachusetts Institute of Technology, 50 Memorial Drive, Cambridge, MA 02139, USA:

BURD RP (1970) Power: An Emerging Subject in Organization Development. SM Thesis, MIT, Cambridge, MA.

COLEMAN DR (1970) A Case Study of Designing, Programming, and Testing a Managerial Visual-Display Model. SM Thesis, MIT, Cambridge, MA.

DEAN JAF (1970) An Analysis of the NDA Review Procedures in the Bureau of Medicine's Office of New Drugs. SM Thesis, MIT, Cambridge, MA.

DIXON GT (1970) Analyzing Student Performance to Improve an Interactive Teaching Program. SM Thesis, MIT, Cambridge, MA.

EBIHARA H (1970) A Study of the World Corporation. SM Thesis, MIT, Cambridge, MA.

FISHER SD (1970) Program Budgeting in Local School Education. SM Thesis, MIT, Cambridge, MA.

GINGOLD JI (1970) The Potential for High Level Computer Manpower in South America. SM Thesis, MIT, Cambridge, MA.

HAMBRECHT GA (1970) National Development Banking Systems and Their Functions. SM Thesis, MIT, Cambridge, MA.

HENDERSON RGC (1970) A Computer Based System for Personal Financial Planning for Middle Income Families. SM Thesis, MIT, Cambridge, MA.

HEYMAN TR (1970) A Model for International and Interregional Market Planning and Control. SM Thesis, MIT, Cambridge, MA.

ISELIN COD JR (1970) East–West Trade and American Technology. SM Thesis, MIT, Cambridge, MA.

JONES R (1975) Administration of Fair Employment Practice Laws by State and Municipal Human Rights Commissions. SM Thesis, MIT, Cambridge, MA.

KINNEBERG DA (1970) Coordinated Bargaining and its Impact on the Collective Bargaining Structure. SM Thesis, MIT, Cambridge, MA.

LAMBRINIDES MA (1970) The Organizational and Psychological Effects of Student Housing on Students. SM Thesis, MIT, Cambridge, MA.

MINER WP (1970) Job Repotting—A Model to Aid in Retaining an Effective and Experienced Management Team. SM Thesis, MIT, Cambridge, MA.

MORGENTHALER GW (1970) Decision Model and Analysis of Corporate Utilization of By-Product R&D. SM Thesis, MIT, Cambridge, MA.

NELSON PW (1970) Effective Routes for Continuing Education Into General Management. SM Thesis, MIT, Cambridge, MA.

PAYETTE RC (1970) Financial Management and Corporate Growth: An Industrial Dynamics Study. SM Thesis, MIT, Cambridge, MA.

PIEPMEIER JM (1970) The Use and Effectiveness of Commiuncations Channels for Scientists and Engineers in a Developing Country—The Concept of the International Technological Gatekeeper. SM Thesis, MIT, Cambridge, MA.

ROMITI VL (1970) EKLOGA (Application of ADP Management Science Techniques in the Analysis of a Normative, Aerospace, New-Oopprtunity Selection Model). SM Thesis, MIT, Cambridge, MA.

ROSENFELD MK (1970) An Input/Output Model for Testing Economics to Scale in the Ethical Pharmaceutical Industry. SM Thesis, MIT, Cambridge, MA.

TANTIVIT B (1970) Effects of Dividend Cuts on Stock Prices. SM Thesis, MIT, Cambridge, MA.

TYMCHUCK DL (1970) The Feasibility of Commercial Introduction of Opaque-2 Corn into the Cali, Colombia Market. SM Thesis, MIT, Cambridge, MA.

VAN DEN BERGH AM (1970) An Inquiry Into the Relationship Between the Mental Processes of a Manager and the Efficiency with which he Fulfills his Task. SM Thesis, MIT, Cambridge, MA.

VANICA PJ (1970) Size, Growth, Performance, and Advisory Fee Arrangements of Registered 'Speculative' Investment Companies, 1960–1969.

ZMUD RW (1970) Internal Pricing for the Multiprogrammed Environment. SM Thesis, MIT, Cambridge, MA.

The following are obtainable at the School of Agricultural Economics and Extension Education, Ontario Agricultural College, University of Guelph N1G 2W1 Canada:

BRADSHAW GJ (1971) Financing Beef Cattle Feedlot Investments: Capital Constraints Confronting Cash-Grain Farmers. MSc Thesis, Ontario Agricultural College.

BROADWITH RA (1972) Explorations of Optimal Feeding Regimes for Growing Pigs. MSc Thesis, Ontario Agricultural College.

CARAMANCION MI (1972) Ontario Market for Livestock Feed. MSc Thesis, Ontario Agricultural College.

CHOI JS (1971) Urban Influence on Agricultural Land Use in the Torronto Centred Region. MSc Thesis, Ontario Agricultural College.

CUSACK TJ (1971) Capital Gains and Losses from Farm Assets in Canada 1946–1967. MSc Thesis, Ontario Agricultural College.

DONKOR F (1971) The Impact of Feedgrain Imports on the Ontario Feedgrain Economy 1946–1968. MSc Thesis, Ontario Agricultural College.

DUKHIA JL (1973) Regional Agricultural Planning in Guyana: A Mathematical Programming Approach. MSc Thesis, Ontario Agricultural College.

HEAD DA (1972) A Study to Determine the Reliability of the Composite-Milk Sampling System in Ontario. MSc Thesis, Ontario Agricultural College.

HSU CY (1970) An Econometric Study of Feedlot Cattle Production. MSc Thesis, Ontario Agricultural College.

KIM YI (1970) Market Potentials for Wheat and Feed Grains in the Far East Countries. MSc Thesis, Ontario A1ricultural College.

KING DLJ (1973) An Economic Appraisal of the Sterile Male Technique of Onion Maggot Control on the Holland Marsh of Ontario. MSc Thesis, Ontario Agricultural College.

LAWRENCE J (1970) Property Taxation in Rural Areas of Ontario. MSc Thesis, Ontario Agricultural College.

LOH MENG K (1971) An Econometric Model of Slaughter and Feeder Cattle Production in Canada and Trade with the United States, 1948–1969. MSc Thesis, Ontario Agricultural College.

MCDONALD A (1970) An Assessment of the Implications of the Ontario Milk Marketing Board's Negotiable Quota Policies with Special Reference to Adjustment within the Fluid Milk Industry. MSc Thesis, Ontario Agricultural College.

MURTI B (1972) Effective Protection and Canadian Sugar Sectors. MSc Thesis, Ontario Agricultural College.

NWACHUKWU SO (1973) A Multi-Period Programming Model for an Economic Evaluation of Cocoa–Oil Palm Farm Settlement Holding in Nigeria. MSc Thesis, Ontario Agricultural College.

Vol. 5
1977

PALLANGYO EP (1972) The Export Prospects for Tanzanian Cotton. MSc Thesis, Ontario Agricultural College.

PANDO JL (1972) The Structural Components of the Supply of Hogs in Canada. MSc Thesis, Ontario Agricultural College.

SINGH H (1971) Implications to Ontario of Changes Taking Place in the US Feed Grain Business. MSc Thesis, Ontario Agricultural College.

SORFLATEN AG (1972) Cost–Output Relationships in Ontario Country Grain Elevators. MSc Thesis, Ontario Agricultural College.

TAN KF (1971) A Case Study of Land Conversion in Grey County. MSc Thesis, Ontario Agricultural College.

TARTE F (1974) A Model of the Feed Purchase Decision Process. MSc Thesis, Ontario Agricultural College.

Information for inclusion in this Section, presented in the appropriate form, should be sent to the Editor of Omega.

Accounting, Organizations and Society

A New International Journal devoted to the Behavioural, Organizational & Social Aspects of Accounting

Edited by:
ANTHONY G. HOPWOOD, *Oxford Centre for Management Studies, Kennington, Oxford*

AIMS & SCOPE

Accounting, Organizations and Society is an international journal devoted to the behavioural, organizational and social aspects of accounting. Published material will range from theoretical and empirical contributions to review articles describing the state of the art in specific areas. Shorter critical assessments of experiments related to the behavioural and social aspects of accounting will also be published. Other features of the journal are book reviews, other bibliographical material and news on related legislative developments. The objectives of the journal are:

1. To provide a specialized forum for the publication of research on the behavioural, organizational and social aspects of accounting.
2. To foster new thinking, research and action on the social and behavioural aspects of accounting.
3. To report on experiments on the behavioural and social aspects of accounting in a way that explains how the experiment was developed, the process by which it was implemented and its consequences, both planned and unplanned. The journal is particularly interested in discussions of the interplay between theory, practice and social and individual values.
4. To accelerate the transfer from research to practice by promoting an international dialogue between researchers, accountants, managers and administrators.

The journal will be concerned with all aspects of the relationship between accounting and human behaviour, organizational structures and processes, and the changing social and political environment of the enterprise. Its unique focus will cover such topics as:

- The social role of accounting, social accounting, social audit and accounting for scarce resources
- the provision of accounting information to employee and trade unions
- processes influencing accounting innovations and the social and political aspects of accounting standard setting
- behavioural studies of the users of accounting information
- information processing views of organizations, and the relationship between accounting and other information systems and organizational structures and processes
- organizational strategies for designing accounting and information systems
- human resource accounting
- accounting and decision-making processes
- the behavioural aspects of budgeting, planning and investment appraisal

SOME ARTICLES TO BE PUBLISHED IN VOLUME 2

Corporate Social Accounting–Reporting for the Physical Environment: A Critical Review and Implementation Proposal: M. DIERKES and L.E. PRESTON

An Inquiry into the Need for Currently Feasible Extensions of the Attest Function in Corporate Annual Reports: W.C. UECKER

Local-Cosmopolitan and Bureaucratic Perceptions in Public Accounting Firms: R.G. SCHROEDER and L.F. IMDIEKE

The Behavioral Sciences and Concepts and Standards for Management Planning and Control: J.G. SAN MIGUEL

A Corporate Strategy for Realizing Equal Employment Opportunity: R.A. VOGT

Role Discrepancies and the Auditor-Client Relationship: K.A. WILCOX and C.H. SMITH

Information Inductance and its Significance for Accounting: P. PRAKASH and A. RAPPAPORT

Perceived Uncertainty and Job Satisfaction in the Accounting Environment: K.R. FERRIS

SUBSCRIPTION INFORMATION

Annual subscription rate (1978)	$ 65.00
Two-year rate (1978/79)	$123.50

Published quarterly. Prices include postage and insurance.

Any individual whose institution takes out a library subscription may purchase a second or additional subscription for personal use at the much reduced rate of $25.00

FREE SPECIMEN COPY AVAILABLE ON REQUEST

Pergamon Press

Pergamon Press Ltd., Headington Hill Hall, Oxford OX3 0BW, England
Pergamon Press Inc., Fairview Park, Elmsford, New York 10523, USA

75

Vol. 5
1977